Defining Dysfunction: Natural Selection, Design, and Drawing a Line*

Peter H. Schwartz†‡

Accounts of the concepts of function and dysfunction have not adequately explained what factors determine the line between low-normal function and dysfunction. I call the challenge of doing so the *line-drawing problem*. Previous approaches emphasize facts involving the action of natural selection (Wakefield 1992a, 1999a, 1999b) or the statistical distribution of levels of functioning in the current population (Boorse 1977, 1997). I point out limitations of these two approaches and present a solution to the line-drawing problem that builds on the second one.

A Case. Mr. Smith is a 70-year-old man being admitted to the hospital with congestive heart failure (CHF). He had a heart attack a few years ago, followed by a procedure that reopened a blocked coronary artery. He did well since then, and tests showed that his "ejection fraction," the amount of blood his heart pushes out in each contraction, remained around 50%, down from the average of 60% but still in the normal range.

Over the last few weeks, he has developed increasing shortness of breath, and tests show that his ejection fraction has dropped to 20%. This reduction in his heart's pumping ability is causing fluid to leak out of the veins in his lungs, causing his trouble breathing. Possible causes of his heart failure range from another heart attack to valvular problems or other issues.

*Received January 2005; revised May 2007.

†To contact the author, please write to: Indiana University Center for Bioethics, 410 W. 10th Street, Suite 3100, Indianapolis, IN 46202; e-mail: phschwar@iupui.edu.

‡Thanks to Christopher Boorse, Randall Dipert, Gary Ebbs, Gary Hatfield, Harold Kincaid, Eric Meslin, Shari Rudavsky, and Alfred Tauber for conversation and encouragement regarding my work on these issues. Thanks to two anonymous reviewers for *Philosophy of Science* for critique and questions that significantly contributed to the final product. Thanks also to audiences at the University of Alabama, Birmingham, the University of Buffalo, and Indiana University–Purdue University, Indianapolis (IUPUI), where I presented earlier versions of this paper.

Philosophy of Science, 74 (July 2007): 364–385. 0031-8248/2007/7403-0003$10.00
Copyright 2007 by the Philosophy of Science Association. All rights reserved.

https://doi.org/10.1086/521970 Published online by Cambridge University Press

1. Introduction. Mr. Smith's heart has gone from pumping blood completely normally, to pumping below average but within the normal range, after his first heart attack, to pumping inadequately. His doctors diagnose him with congestive heart failure, and he now counts as having a disease. Looked at biologically, his heart is dysfunctional since it is failing to carry out its proper function of pumping blood.[1]

Mr. Smith's heart has crossed a line, from being normal, healthy, and properly functioning, to being abnormal, diseased, and dysfunctional. I will argue that the line between properly functioning and dysfunctioning is poorly understood, and I will call the challenge of providing an adequate account of this distinction the *line-drawing problem*. The problem is important, first, since it raises an issue that has been overlooked in the debate over "function" and teleology more generally in philosophy of biology. Second, in medical ethics and philosophy of medicine, the problem poses a challenge to certain important definitions of "disease," which rely on criteria involving dysfunction.

I'll start, in Section 2, by explaining why the line-drawing problem is relevant to analyzing the biological concept of function in biology. In Section 3 through Section 5, I examine two definitions of "disease," presented by Christopher Boorse (1977, 1987, 1997) and Jerome Wakefield (1992a, 1992b, 1999a, 1999b), which both make dysfunction necessary to the presence of disease in an organism. Although these theorists have not formulated the line-drawing problem as I do here, in the process of spelling out their accounts each has suggested a possible approach. I will examine their proposals and argue that they fall short in instructive ways. In Section 6, I propose my own approach, building on Boorse's.

2. Function and Line Drawing. The concept of dysfunction has been generally ignored by philosophers writing about function (two exceptions are Neander [1995] and Boorse [1977, 1997]). Theorists have focused, instead, on the question of which of an item's many effects or possible effects count as its function. The heart, for instance, pumps blood, makes noise, generates heat, and burns glucose, but only pumping blood counts as its function. The worry has been that picking out some effect as the *function* involves problematic teleological ideas or occurs as part of illegitimate attempts at teleological explanation.

Philosophers have responded by proposing definitions of "function"

1. For the purposes of this paper, I will treat the terms "dysfunctional" and "dysfunctioning" interchangeably. And while I believe these terms are also interchangeable with "malfunctional" and "malfunctioning," I will avoid the latter two for simplicity's sake. There may be those who believe there are important and systematic differences between these terms, but I believe this is questionable.

https://doi.org/10.1086/521970 Published online by Cambridge University Press

based on scientific, nonteleological notions and by describing types of explanations where function ascriptions can play legitimate roles. One major approach in philosophy of biology has been the "etiological view," where an item's proper function is (roughly) the effect that was favored by natural selection (e.g., Millikan 1989; Neander 1991). Nonetiological accounts, in contrast, argue that there is no need to invoke natural selection and instead utilize other causal and statistical notions (e.g., Boorse 1976; Cummins 1975).

None of these theories have focused on the line-drawing problem. But drawing the line between normal and abnormal levels of pumping blood, for instance, raises the same teleological worries as the question of why pumping, rather than other effects of the heart, counts as its function. Just as highlighting one of the heart's effects as its function raises the worry that a judgment is being made about what the heart "should do" (or about its "purpose," etc.), so does picking out one level of pumping as normal. A heart that beats with an EF of 50% is doing what hearts "should do," while one that beats with an EF of 20% is failing to do so. Once again, teleological ideas may be creeping into the description of the heart.

The same questions arise in analyzing the central biological idea of "normal variation." A generally accepted principle of modern biology is that there is no single design for any species, but instead a normal range of variation for almost every trait. Even among healthy 70-year-old men, for instance, there is a large range in the ability of hearts to pump. People who run marathons have very efficient hearts, reflected in their EF and other features, while those who rarely exercise have much less efficient hearts. But there are boundaries to this normal range: below a certain level of functioning, the heart counts as abnormal. Making sense of "normal variation" requires solving the line-drawing problem as well.

3. Dysfunction-Requiring Accounts of Disease.

3.1. Advantages of Dysfunction-Requiring Accounts. Of the function theorists, Christopher Boorse (1977, 1987, 1997) has confronted the line-drawing problem most directly, in his attempt to use the notion of dysfunction to present a value-free definition of "disease." Jerome Wakefield (1992a, 1992b, 1999a, 1999b) has more recently presented a definition of "disease" that also includes a criterion involving dysfunction, and this leads him to suggest an approach to the line-drawing problem as well, although a less developed one than Boorse's.

The "dysfunction-requiring" (DR) theories of disease, as I will call them, form a small subset of the many definitions that have been put forward, and the DR view is interesting partly because it avoids serious

https://doi.org/10.1086/521970 Published online by Cambridge University Press

problems that other approaches face. For instance, Culver and Gert (1982, 81), define a "malady" as a condition that causes a person to be "suffering, or at increased risk of suffering, an evil (death, pain, disability, loss of freedom or opportunity, or loss of pleasure) in the absence of a distinct sustaining cause". But this definition would make pregnancy a malady, as the authors later accept (Gert, et al. 1997, 126). Also, since certain desires or beliefs may increase one's risk of suffering an "evil"—e.g., perhaps a desire to climb a mountain, or a belief that the government in power is up to no good—Culver and Gert (1982) include a specific condition ruling out rational beliefs or desires from being maladies. But this way of solving the problem seems horribly ad hoc.

DR theories of disease count none of these conditions—wanting to climb a mountain, being pregnant, distrusting the government, etc.—as diseases since none of them involve biological dysfunction.[2] It is partly because of these advantages that Boorse's account serves as the basis for one of the best worked-out accounts of the ethics of healthcare distribution (Daniels 1985; see also Buchanan et al. 2000, Chapters 3–4).

3.2. Two Theories. Here are the two DR accounts I will discuss. Boorse calls his approach the "Biostatistical Theory" and states it recently as follows:

1. The *reference class* is a natural class of organisms of uniform functional design; specifically, an age group of a sex of a species.
2. A *normal function* of a part or process within members of the reference class is a statistically typical contribution by it to their individual survival and reproduction.
3. A *disease* is a type of internal state which is either an impairment of normal functional ability, i.e., a reduction of one or more functional abilities below typical efficiency, or a limitation on functional ability caused by environmental agents.
4. *Health* is the absence of disease. (Boorse 1997, 7–8)

Wakefield presents his "Harmful Dysfunction Account" (HD) as follows:

> A condition is a disorder if and only if (a) the condition causes some harm or deprivation of benefit to the person as judged by the standards of the person's culture (the value criterion), and (b) the condition results from the inability of some internal mechanism to per-

2. DR theories, of course, have problems of their own, and I cannot attempt to address them here. For some of the strongest defenses, see Boorse (1997) and Wakefield (1992a, 1999a, 1999b).

https://doi.org/10.1086/521970 Published online by Cambridge University Press

> form its natural function, wherein a natural function is an effect that is part of the evolutionary explanation of the existence and structure of the mechanism (the explanatory criterion). (1992a, 384)

There are a few similarities and differences between the two theories to note immediately. First, they vary in their choice of terms—Boorse uses "disease" and Wakefield "disorder"—but both writers specify that they are aiming at all pathological conditions (Boorse 1997, 11; Wakefield 1999a, 376). I'll concentrate on cases of physical illness, and I'll switch between the three terms—"disease," "disorder," and "pathology"—freely. Second, while Wakefield considers the presence of dysfunction to be a necessary condition for the presence of disease, Boorse treats it as both necessary and sufficient. This difference will not matter to the discussion here, since reliance on a notion of dysfunction in either way raises the line-drawing problem.

Third, Boorse and Wakefield base their definitions of "dysfunction" on different accounts of function, since Boorse's account (in Premise 2) requires just a typical contribution to survival and reproduction, while Wakefield requires that the trait was favored by natural selection. This difference will not matter to my discussion of cases like Mr. Smith's heart failure, in the analysis of the line-drawing problem, since both approaches assign the human heart the function of pumping blood.

Finally, Boorse and Wakefield both offer their accounts as *conceptual analysis*. Because of this orientation, each formulates his definition as a short list of necessary and sufficient criteria, and each argues that his definition fits with medical opinion and other intuitions about what is a disease, using a wide range of real and imagined cases. There are various problems with this way of framing the debate (cf. Ramsey 1992), but I'll play by the same rules here. For further discussion, see Section 7, below.

It's important to note also that Boorse and Wakefield both reject any quick inference from the disease-status of a condition to any conclusion about whether it should be treated or not (e.g., Boorse 1997, 11–13; Wakefield 1999a, 374). While both view medicine as having a particular interest in treating and preventing diseases, they agree it may be best to treat some nondiseases and not treat some diseases. Following them and others, I will focus on the analysis of concepts of disease and dysfunction, avoiding conclusions about the advisability of medical interventions.[3]

3. One view that treats the disease-status of a condition as relevant to its treatment by medical professionals, without making disease-status necessary or sufficient for such treatment, is the "Primary Rationale for Medical Obligations" view (Buchanan et al. 2000, 121–124).

https://doi.org/10.1086/521970 Published online by Cambridge University Press

4. Wakefield's Approach.

4.1. Wakefield's Approach to the Line-Drawing Problem. The essence of Wakefield's approach to the line-drawing problem involves his "explanatory criterion," which requires that "the condition results from the inability of some internal mechanism to perform its natural function" (1992a, 384). The simplest way to interpret this would be as requiring true *inability*, i.e., that the organ has no capacity to carry out its function at all. But that can't be right, since Mr. Smith's heart is dysfunctioning, and yet it is still pumping, just at a radically decreased level.

Wakefield needs to provide some other account of what "inability" means. He makes a number of comments on this topic, mostly invoking the idea of design. He writes, "Roughly, we recognize a variation in anatomical structure as a lesion rather than as a normal variation if the variation impairs the ability of the particular structure to accomplish the functions *that it was designed to perform*" (1992a, 375, emphasis added). In another paper, he writes, "A dysfunction exists when a person's internal mechanisms are not able to function in the range of environments for which they were designed" (1992b, 243).

He makes it clear that he means "design" in these contexts as referring to the action of natural selection. He writes that dysfunction is ". . . 'failure of a mechanism in the person to perform a natural function for which the mechanism was designed by natural selection'" (1992b, 236). Later, he writes, ". . . it is standard among evolutionary theorists to use design, purpose, goal, and other intentional metaphors to describe the results of natural selection" (1999a, 376).

This last comment underestimates the magnitude of the controversy over the use of such metaphors in biology. On one side, some philosophers of biology argue that because living things are not actually "designed," the term can apply only metaphorically (e.g., Lewontin 1983; Matthen 1997). Others hold that organisms are *literally* designed by natural selection, although in a different sense of "design" than when the term refers to manmade artifacts (e.g., Allen and Bekoff 1995).[4] We can sidestep this dispute by noting that on either view, if "design" means anything substantial in a biological context, it refers in some way to natural selection.[5] And for both views the absence of an intentional designer in natural selection blocks any simple inference from "natural design" to "intentional design."

4. It's not easy to classify Dennett (1995) as falling in either of these camps: although he extensively defends the idea that Natural Selection designs living things, he commonly puts "design" in scare quotes in this context.

5. Lewens (2004, especially Chapter 1) makes a similar observation.

https://doi.org/10.1086/521970 Published online by Cambridge University Press

In his comments that apply directly to the line-drawing problem, Wakefield writes that natural selection favors a "range of responses" in organisms (1999a, 379), and he argues that "a range of selected values can be distinguished from nonselected values and this is all that is required for the HD analysis to discriminate disorder from nondisorder" (1999a, 387).[6] So, the existence of wide variation in functioning, such as in hearts' pumping ability, does not pose a problem for his theory, he writes, since only some levels were favored by selection, while others were selected against. Discussing the example of panic attacks, he writes, "there is a naturally selected range of the sensitivity of fear-response mechanisms, but the spontaneous terrors of panic disorder are not part of that range" (1999a, 387).

4.2. Problems for Wakefield's Approach. But this approach does not solve the line-drawing problem. The first difficulty is that for acquired disorders, ones not caused genetically, natural selection does not apply. For example, it is not the case that Mr. Smith's EF of 20% is favored by natural selection or not favored, since his condition is not hereditary. Maybe Wakefield would recommend that we identify some cutoff for "heart failure" by drawing an analogy to hereditary heart problems, such as congenital heart malformations. Perhaps there is some ejection fraction that is low enough that we can confidently conclude that natural selection would discriminate against individuals born with that EF.

But this approach leads to the second, and deeper, problem with Wakefield's proposal. Deciding whether a certain variant is *selected against* depends on which other forms exist in the population. If a certain hereditary condition causes children at birth to have an EF of 20%, and only 1% of these children will survive to reproduce, this genetic variant will still be favored by selection compared to conditions carrying a 0.5% chance of survival. Consider the following imaginary case of five variants with their associated levels of expected reproduction: Which ones count as being "selected for" and which are "selected against"?

Variant	Expected Reproduction
A	.2
B	.5
C	1
D	2
E	5

6. This in response to criticism from Lilienfeld and Marino (1995, 414), who point out that there is no single level of functioning that is favored by natural selection. My point here is meant to build on theirs.

https://doi.org/10.1086/521970 Published online by Cambridge University Press

Variants B–E are selectively *favored* in comparison to A, and variants A–D are *selected against* in comparison to E. It's tempting to count just variant A as the pathological one, since it has the lowest expected reproduction, but variant B also has very low fitness and thus might be taken as falling outside the normal range. And if we count B as outside the normal range, we might want to count C as selected against as well. Finally, if we discover another variant that is incompatible with life (expected reproduction = 0), does this mean that even variant A falls in the normal range?

Similar reasoning applies to the case of panic disorder: there may be a range in "the sensitivity of fear-response mechanisms," as Wakefield says (1999a, 387), and if these are hereditary there may be a fitness associated with each. But there is no line here that can discriminate between which ones are naturally selected and which are not, despite Wakefield's confidence (1999a, 387). Something other than the existence of positive or negative selection must be used to solve the line-drawing problem.[7]

4.3. Design and Artifacts. Determining whether a manmade artifact is dysfunctioning often depends on facts about the designer's intentions, and it is the absence of such intentions in the case of organisms "designed" by natural selection that sinks Wakefield's approach to the line-drawing problem. For example, consider my laser printer. Laser printers on the market today display a great variety in resolution, printing speed, durability, and ability to print on papers of different sizes and weights. Since my printer is a low-cost one meant for home offices, it does relatively poorly on all these measures. But for now it is in perfect working order, since it is operating exactly as planned.

Even among printers of the same make and model, there will be some variation in performance due to factors such as how well certain parts were welded together or lubricated. For example, imagine that printers of my make and model coming out of the factory may print anywhere from 12 pages per minute (ppm) to eight pages per minute. What determines which levels of printing are normal—and what tolerance among the printer's parts are acceptable—involves the printer's specifications (set by the manufacturer), the promises made in advertising, and the understanding of the general purchaser. These are exactly the sort of factors that are completely absence in the case of "design" by natural selection. Although the analogy between design by humans and by natural selection may be helpful in some ways, relying on it to solve the line-drawing problem cannot work.

7. Boorse (2002, 101–102) makes a related point.

https://doi.org/10.1086/521970 Published online by Cambridge University Press

5. Boorse's Proposal.

5.1. Premises 1 and 2. Boorse's articles (1977, 1987, 1997) suggest a different approach to the line-drawing problem, although he doesn't focus on the issue either. Consider his four-part definition of "disease," above (Section 3.1), starting with his idea of a "reference class" in Premise 1. According to this premise, a trait-token's level of functioning is compared against the trait-tokens of other individuals of the same gender and age. The discussion of Mr. Smith's heart failure, above (Sections 2 and 4.2), already implicitly assumed this, since his heart's pumping ability was compared to that of other 70-year-old men, not of young adults or women, for instance.

Reference classes are necessary because of the variability among age groups and between the two genders. In humans, the female breast has the function of producing milk at certain times, for example, while the male breast never does. Human children younger than 6 months old can't walk and newly hatched chicks can't fly, but their legs and wings (respectively) do not count as dysfunctional. As individuals age, the level of functioning considered normal declines: a certain EF may count as normal in an 80-year-old when it would be abnormal in a 20-year-old. Even cessation of function may be normal, like menopause in a 50-year-old woman. Menopause at younger ages, termed "premature ovarian failure," is not normal.

Boorse states his theory of function in Premise 2. One small modification in his theory is necessary to allow it to apply to older individuals, such as Mr. Smith. According to Premise 2, the hearts of 70-year-old men only have the function of pumping blood if that is the "statistically typical contribution" of these trait-tokens to "individual survival and reproduction" (1997, 7). But since men at this age generally have finished reproducing, it may turn out that hearts in this reference class make no contribution to *reproduction* and thus, formally, make no contribution to "survival *and* reproduction." Simply changing the conjunction to disjunction resolves this problem: a trait-type's function is its statistically typical contribution to "survival *or* reproduction."[8]

5.2. The Frequency Approach. Premise 3 presents Boorse's definition of disease, treating the term as basically synonymous with "dysfunction." So the key to answering the line-drawing problem for Boorse is deciding

8. Note that while this change helps Boorse's theory here, it may raise problems for his account of functions in other areas by making it even more permissive than it already is. For example, a trait that extends survival slightly by interfering with successful reproduction may be counted as having the function of doing so.

https://doi.org/10.1086/521970 Published online by Cambridge University Press

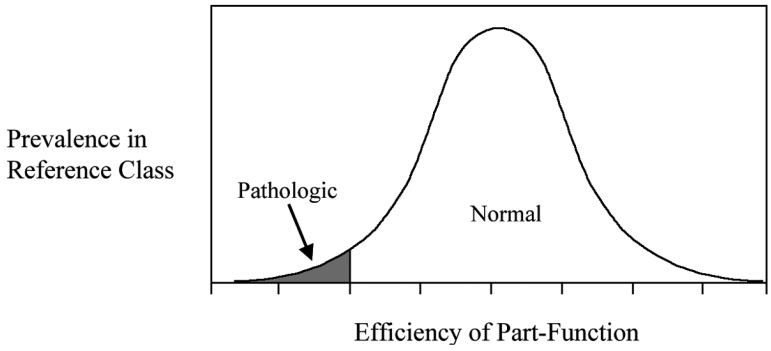


Figure 1. Adapted from Boorse (1987, 370; 1997, 8). (I've reversed the axes, to place them as Boorse initially intended [personal communication] and in keeping with the standard practice of placing the dependent variable on the *Y*-axis.)

when there is "an impairment of normal functional ability, i.e., a reduction of one or more functional abilities below typical efficiency" (Boorse 1997, 7–8).[9]

One possible interpretation of "below typical efficiency" will definitely not work, i.e., taking it as anything falling below mean (or median) efficiency. This would make all organs functioning below average count as dysfunctional and that can't be right. As Boorse says, a condition is a dysfunction only if it "falls *more than a certain distance below* the population mean" (1977, 559; emphasis added). And this raises the line-drawing problem again: How far is far enough?

The little Boorse says about this topic centers on a graph of "Efficiency of part-function" vs. "Statistical distribution in reference class" (1987, 370; 1997, 8), reproduced with some modifications in Figure 1. He assumes that the distribution is statistically normal and writes that the cutoff between low-normal function and dysfunction "can only be conventionally chosen, as in any application of statistical normality to a continuous distribution. The precise line between health and disease is usually academic, since most diseases involve functional deficits that are unusual by any reasonable standard" (1977, 559). In a later paper, he writes, "[T]he lower limit of normal functional ability—the line between normal and pathological—is arbitrary. Although statisticians often use . . . [a] 95 percent central range, no reason for such a choice applies here. The con-

9. The last phrase of Premise 3, which says that a trait is dysfunctional if it is "a limitation on functional ability caused by environmental agents," (Boorse 1997, 8) will only be relevant for cases of unusually prevalent diseases, discussed below (Section 5.3, FN's 12 and 13).

https://doi.org/10.1086/521970 Published online by Cambridge University Press

cept of a pathological state has vague boundaries—though the vast majority of disease processes involve functional deficits by any reasonable standard" (1987, 371). Interpreting these comments, and thus Boorse's approach to the line-drawing problem, requires some care.

First, Boorse writes that the line is "conventionally chosen" or "arbitrary."[10] That said, he clearly does not mean that the line could be drawn anywhere, since this would imply relativism about disease-status, which Boorse clearly rejects. For example, even if we want our children to be taller, it would be a mistake to say that a child in the sixtieth percentile for height suffers from "dwarfism." Only certain places to draw the edge of normality are acceptable.

Second, Boorse mentions the existence of "functional deficits" as confirming the judgment that a condition is a disease. But such deficits are not necessary, as we can see from his comment that just a "majority" or "vast majority" of diseases have them. And if there are deficits, they need not be noticeable at a tissue or organismic level. He writes: "Local part-dysfunctions need not have any gross effects on disability or deformity or distress. . . . Liver cells, to be normal, must perform a host of metabolic functions because that is what liver cells collectively contribute to survival and reproduction. But a large number of liver cells can be pathological without clinically detectable effects or appreciable risk of such effects" (1987, 371–372). In fact, one of the strengths of Boorse's theory is that it makes disease-status independent of facts about whether the condition carries undesirable consequences for the organism. This independence helps the theory avoid the problems that plague non-DR theories such as those emphasizing harm (see Section 3.1).

Finally, although Boorse says that there is nothing inviolate about the "95 percent central range" around average, he does give statistical measures primacy. He writes, "health is normal functioning, where the normality is statistical and the functions biological" (1977, 542). He emphasizes the importance of statistics when he calls his account the "Biostatistical theory" and says that "disease is only statistically species-subnormal biological part-function" (1997, 4). Note also that the graph in Figure 1 just represents the prevalence of each level of function, without providing any information about the consequences of each level.

I'll call the view that Boorse appears to embrace the *Frequency approach* to the line-drawing problem. According to this approach, the range of acceptable places to draw a line between low-normal and dysfunction is determined by statistical measures of the distribution of levels of functioning in the reference class. For example, if there is a normal distribution

10. Wakefield says something similar, writing that for concepts such as disorder, "boundaries are set partly by convention" (1999a, 379).

https://doi.org/10.1086/521970 Published online by Cambridge University Press

of levels, then the line must be drawn near two standard deviations below the mean. The exact location can be chosen arbitrarily as long as it falls in this area.

5.3. The Problem of Common Disease. The first challenge for the frequency approach can be described again using Mr. Smith's case of heart failure. The relevant graph in this case is Ejection Fraction (EF) on the *X*-axis and prevalence among 70-year-old men on the *Y*-axis. Assuming that Mr. Smith's EF of 20% falls more than two standard deviations below the mean—placing him in the bottom 1% of members of his reference class, say—then the Frequency approach properly classifies his heart as dysfunctional. But the approach also implies that if an EF of 20% were more prevalent, say found in 10% or 20% of the population, then it would not count as dysfunction or disease. And this would be so even if all these people felt the same negative consequences as Mr. Smith. I'll call this the *problem of common disease*.

At first blush, this might sound similar to some well-known critiques of Boorse that use examples of common disease, which emphasize that a nonetiological account of function like his has trouble explaining how a trait-type continues to have the function F in situations where all (or almost all) trait-tokens fail to do F (Millikan 1989, 300; Neander 1991, 182–183). The challenge posed by the problem of common disease is different, however, for a number of reasons. First, it arises in cases where function is reduced, not lost. Second, the problem arises even if only a significant minority of individuals in the reference class (e.g., 20% or 30%) have decreased function. Third, the problem of common diseases faces both etiological and nonetiological accounts of function, i.e., any theory that adopts the frequency approach for distinguishing between normal and abnormal.

Finally and perhaps most importantly, there are no actual cases of universal disease (as far as I'm aware), and there are plenty of real-life cases of common disease. For example, in dogs the prevalence of serious dysfunction of the hip joint, termed "canine hip dysplasia" has been estimated at more than 30% in some breeds (Rettenmaier et al. 2002). In humans, there are many prevalent dysfunctions, especially in the elderly. For example, serious urinary dysfunction due to Benign Prostatic Hypertrophy (BPH) occurs in more than 17% of men over 70 (Bosch et al. 1995),[11] and senility of the Alzheimer's type has a prevalence of 16% in people over 85 years old (Polvikoski et al. 2001).

11. These numbers for prevalence are calculated using some of the most restrictive possible definitions of BPH. Less demanding definitions yield prevalences as high as 28% in this age group (Bosch et al. 1995, 38).

https://doi.org/10.1086/521970 Published online by Cambridge University Press

At the end of Premise 3, Boorse adds a codicil that allows a prevalent condition to count as a disease, i.e., if it is "a limitation on functional ability caused by environmental agents" (1997, 8). This criterion is meant to handle cases of universal disease, where a virus or toxin becomes common.[12] But it will not help explain why BPH and Alzheimer's, or hip dysplasia in dogs, count as diseases, since there is no clear environmental cause for their prevalence.[13]

5.4. The Problem of Healthy Populations. The second major problem facing the Frequency approach arises if we think about reference classes where dysfunction is incredibly rare. Imagine, for instance, that EF in 20-year-old men ranged just from 50% to 70%, and physicians thus did not diagnose any men at this age with congestive heart failure. In such a case, the Frequency approach implies that as long as there is a normal distribution of functional abilities, the lowest 1–2% must count as dysfunctional. The *problem of healthy populations*, as I'll call it, arises in any case where a trait's lowest level of functioning in some reference class is not low enough to be dysfunctional.

Wakefield's definition of "disorder" avoids the problem since his first criterion says that a condition only counts as a disorder if it causes harm. Thus, in our imaginary population of healthy 20-year-olds, the ones with an EF below the first percentile in pumping efficiency won't count as having a disorder since their lower-than-average EF causes them no harm. That said, the puzzle remains why they should count as having a dysfunction at all, i.e., as fulfilling Wakefield's second criterion for disorder. And if Wakefield were to adopt the Frequency approach to defining dysfunction, he would still have to address the problem of common disease.

6. The Frequency and Negative Consequences (FNC) Approach.

6.1. A Proposal. The problem of common disease and the problem of healthy populations show that there is more variability in the prevalence of dysfunction than the Frequency approach allows. What appears to be needed is an additional factor, and a natural candidate is the effect that a given level of functioning has on the organism. For example, in the problem of healthy populations, it is the lack of negative consequences

12. Note that the condition does nothing to explain why the trait-type in such cases should count as still having the function F at all, since it no longer makes any contribution to survival or reproduction. Answering that problem requires interpretation or modification of Premise 2.

13. Boorse (2002, 103) has recently discussed dropping this additional clause from Premise 3, due to a variety of challenges it faces.

https://doi.org/10.1086/521970 Published online by Cambridge University Press

of having an EF of 50% that makes this level of pumping appear to be low-normal, rather than dysfunctional, even though it falls in the bottom 1% of the reference class. In the problem of common disease, the presence of important negative consequences is what makes serious BPH or Alzheimer's appear to be dysfunctional, even though the conditions are so prevalent.

One way to introduce a role for facts about negative consequences is to add a third dimension to the sort of graphs used in the Frequency approach. In addition to the *X*-axis showing the level of functioning and the *Y*-axis showing prevalence, one can add a *Z*-axis representing the magnitude of negative consequences imposed by the functional level. The role for the magnitudes displayed on the *Z*-axis would be the following. More severe negative consequences would support moving the line between normal and abnormal (on the *X*-axis) toward the *right*, to count more levels of functioning as dysfunctional. A lack of negative consequences would support moving the line to the *left*, to count only lower levels of functioning (or even no levels of functioning at all) as dysfunctional. This method presents no simple rule about where to put the line, but it provide a role for judgments about consequences, and thus a way to answer the problems facing the Frequency approach.[14] Call this the Frequency and Negative Consequences (FNC) approach.

Consider the problem of common disease, such as BPH in 70-year-old men or Alzheimer's in 85-year-old men or women. In these cases, the negative consequences would support moving the line between low-normal and dysfunctional on the *X*-axis to the right. Similarly, the serious effect of canine hip-dysplasia on ambulation would support considering the lowest 30% of the breed as having dysfunctional hips. In the sort of case that arises in the problem of healthy populations, the lack of negative consequences of having an EF of 50% would support moving the line between normal and abnormal to the left on the *X*-axis. Thus an EF of 50% might well count as normal even though its prevalence is 1% in the reference class.

The FNC approach is illustrated by the graphs in Figure 2a and Figure 2b, with Figure 2a showing the case where 30% of individuals in a dog species has dysfunctional hips, and Figure 2b showing a case of healthy populations where there are no cases of CHF. (Note that due to the difficulty representing a three-dimensional figure here, I graph the negative consequences of a given level of functioning as a second *Y*-axis. Prevalence of a level of functioning is shown by the solid line, with magnitudes labeled

14. I am grateful to one of the anonymous referees of *Philosophy of Science*, who made a suggestion that was crucial to my choosing this way of combining facts about negative consequences and prevalence.

https://doi.org/10.1086/521970 Published online by Cambridge University Press

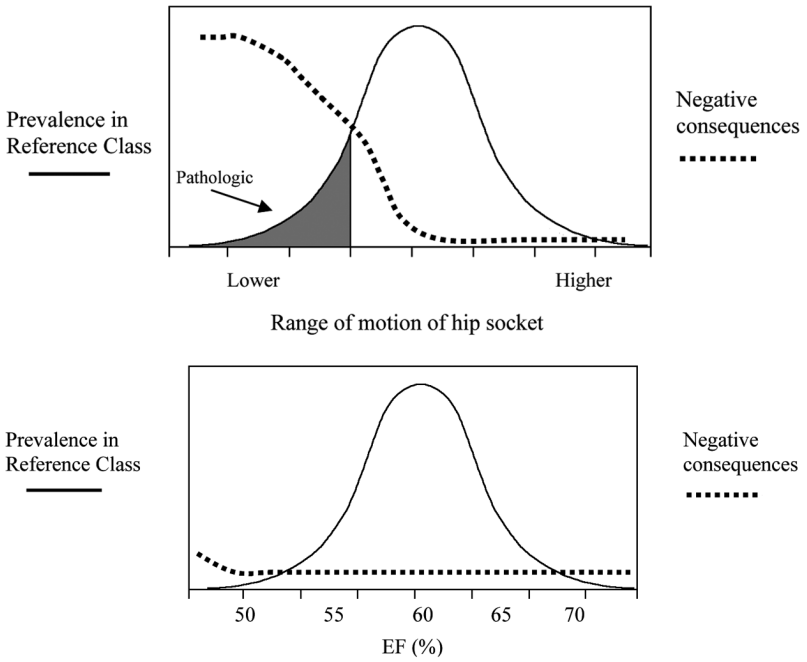


Figure 2. a. Prevalence and negative consequences of canine hip functioning. b. Prevalance and negative consequences of EF in healthy population.

on the left side of the graph, while severity of negative consequences is indicated by a dotted line, with magnitudes labeled on the right side of the graph.)

To flesh out the way that negative consequences come into play in cases like this, consider the following thought experiment. Imagine that there are two areas of the human brain—areas "A1" and "A2"—that are involved in motor coordination. Imagine that all people lose cells in both areas every year, with some people losing more and some less, and the capacity of A1 or A2 is directly proportional to the number of cells currently present in that area at a given time. Assume that in any age group, there is a statistically normal distribution in the number of cells and thus the capabilities of the area. Finally, imagine that in the reference class of 70-year-old men, those whose A1 functions in the lowest 30% have symptoms of moderate Parkinson's disease, while only those whose A2 functions in the lowest 2% show such symptoms.[15]

15. The loss of cells in the substantia nigra during the development of Parkinson's disease is roughly similar to the sort of process I am imagining for the development of pathology in areas A1 and A2.

https://doi.org/10.1086/521970 Published online by Cambridge University Press

It seems clear that the lowest functioning 30% of A1's will count as dysfunctional, while just the lowest 2% of A2's will be so classified. Biologists—and doctors and medical scientists, in this case, since the species is *Homo sapiens*—would respond to the presence of the symptoms by looking for a cause. When it is found to be the low level of functioning in A1 or A2, these levels would be judged dysfunctional, i.e., the lowest 30% in A1 and the lowest 2% in A2. For individuals with an A2 functioning at a level above 2%, there would be no temptation to count the area as dysfunctioning.

6.2. Specifications. Making FNC work requires specifying carefully how consequences are evaluated and quantified. One approach that definitely does not work is to count only reductions in ability to survive and reproduce. Interestingly, some critics mistakenly attribute a position like this to Boorse, overlooking his comments favoring the Frequency approach. Wakefield, for example, writes that for Boorse, "a disorder is a condition that reduces longevity or fertility" (1992a, 378). Although these critics are wrong to attribute the position to Boorse, they're right to attack it. The main problem is that dysfunctions don't necessarily reduce survival or reproduction at all. If an individual has aphasia, i.e., an inability to speak (Wakefield's example [1999a, 379]), or blindness, or, for that matter, is missing a limb, he won't necessarily have lowered expected survival or reproduction if he lives in a society that is suitably supportive. As Boorse points out, in certain settings a disease may even improve survival and reproduction, as when flat feet keep a recruit out of the army (1977, 545). Any measure of "negative consequences" needs to be more fine-grained and contextual than just looking at effects on survival and reproduction.

We also have to avoid making "negative consequences" too broad. For example, if negative consequences were entirely in the eyes of the beholder—including *any* effect that has negative valence to the individual—then the problem of healthy populations would immediately be regenerated for FNC. If some individual with EF 50% cares deeply about having a higher numerical value of his EF—due to vanity, for instance, or due to a desire to carry out some unusual activity that it would allow—then that individual would again count as having a dysfunctioning heart.

Instead, the relevant negative consequences should be those that impact some standard activity or capacity of the organism. More precisely, they should be effects that significantly diminish the ability of a part or process in the organism, or of the overall organism, to carry out an activity that is generally standard in the species and has been for a long period of time. In most cases, and perhaps all (I cannot settle this fine point here), this will be an activity or capacity that has been subject to the process of natural selection in the species. Classic cases of dysfunction—such as

https://doi.org/10.1086/521970 Published online by Cambridge University Press

aphasia, blindness, deafness, or flat feet—all represent significant decreases in the organism-level capacities of speaking, seeing, hearing, and running, respectively, whether or not there is any effect on survival and reproduction. A person who is blind due to chemical damage to his corneas for instance, has a level of functioning of the visual system that falls significantly below that which has been standard in humans for eons, and was certainly the product of natural selection acting on humans and their ancestors.

6.3. Problems of Including Too Much. With the FNC account specified in this way, possible attacks can come from two directions, i.e., charging that it includes too many conditions as dysfunctions or too few. An attack of the first sort might claim that FNC account counts some undesirable but normal conditions as dysfunctions. Take normal aging, for instance. According to the Frequency approach, the standard changes of aging are not dysfunctions, since they are so common. But since the FNC account factors in negative consequences—making it possible that common states can count as dysfunctional—there is at least the danger that the undesirable effects of aging could cause the line between low-normal and dysfunction to slide far enough to the right to classify some of normal aging as dysfunction.

This attack makes most sense as a claim that the FNC account will count the normal *process* of aging, rather than an *effect*, as dysfunctional, since although the process of aging is generally considered normal, the process can lead to the creation of dysfunctional states. For example, if a person has an already reduced EF—due to previous damage, such as in Mr. Smith's initial heart attack—it is possible that normal aging could reduce the EF further and set off congestive heart failure. For this reason, an attack on the FNC account should focus on the danger that this approach will classify the normal aging *process* as dysfunctional.

But the attack is not concerning, for a number of reasons. First, in some people the process of aging may reflect dysfunction, e.g., if it is unusually rapid, as in the disease progeria. So classifying the process as dysfunctional in some people may not be mistaken. Second, even though typical aging may be "natural" in some sense, current evolutionary theories suggest that it does not reflect the proper functioning of any part or process in the body (Caplan 2004), and thus there is no danger that the process in a particular person could represent dysfunction. According to these theories, natural selection favored processes that allow individuals to reproduce rapidly, with these same processes causing senescence when they continue after reproduction is over. Thus, aging after reproduction would have no function according to etiological or nonetiological accounts. In fact, these biological theories have led some to argue that

https://doi.org/10.1086/521970 Published online by Cambridge University Press

"normal" aging is a form of *disease*, albeit a universal one (Caplan 2004). For all these reasons, an attack on the FNC account involving its interpretation of normal aging should not give us pause.

6.4. Problems with Including Too Little. The second type of attack on the FNC approach could claim that its requirement that there be negative consequences leads to its classifying *too few* conditions as dysfunctional. The common cold or hay fever, for instance, are well-recognized diseases, and thus, according to DR accounts, they must involve dysfunction. But their negative consequences may be thought to be too minimal to satisfy the FNC's requirement.[16]

I believe that the FNC can easily handle such cases. First, the negative consequences of these conditions on respiration (and the normal functioning of nasal mucosa, for instance) are significant enough to satisfy FNC's requirements. A stuffy nose may not cause a reduction in survival and reproduction, but Section 6.2 (above) showed that such effects are not necessary in order for consequences to count as significant. People with serious hayfever may have a greatly reduced ability to moisen and warm air being inhaled through their nose, for instance.

Second, even if the consequences were quite minor, it would be more of a problem for the DR approach to disease than for the FNC account of dysfunction. Many mild conditions commonly classified as diseases do not involve any clear dysfunction, a problem that has been raised for DR accounts (Schwartz 2007). Lastly, given the prevalence of hay fever in the American population currently, this condition may pose a more serious attack on a DR account of disease adopting the Frequency approach than one adopting the FNC approach.

A greater challenge to the FNC account stems from cases where a clear dysfunction has no negative consequences at all, i.e., truly asymptomatic conditions. Both Boorse and Wakefield discuss such conditions, defending their focus on dysfunction rather than negative consequences in defining disease.[17] Boorse mentions that large numbers of liver cells can be dysfunctional without there being any impact on the liver's overall function (1987, 371–372, quoted in Section 5.2, above). Wakefield discusses a case where a small number of sperm are dysfunctional, unable to swim towards the egg, although there are enough other functional sperm so that there is no overall decrease in fertility (Wakefield 1999a, 376).

Note that these cases are similar in that they are both ones where there

16. I am thankful to an anonymous referee at *Philosophy of Science* for raising this concern.

17. Once again, I am thankful to an anonymous referee at *Philosophy of Science* for emphasizing the importance of addressing this sort of case.

https://doi.org/10.1086/521970 Published online by Cambridge University Press

is such excess capacity that that the dysfunction of a part (or a number of cells) has no noticeable impact on the whole. I will argue that such cases are properly understood as ones where the overall organ or group of cells is functioning properly, even if a part is dysfunctioning, and I will show that the FNC approach can yield this result.

Focusing on Boorse's case of the asymptomatic liver condition, three different ways of specifying the case appear to make an important difference in its implications. First imagine that a number of liver cells have died, as part of a process that is completely standard in humans, perhaps due to apoptosis. In this case, there's no question whether these cells are functioning or dysfunctioning, since they are dead, and there is little attraction to thinking that the overall liver is functioning abnormally, since the loss of these cells is typical. The FNC account would concur.

Second, imagine that instead of the cells dying, they have just lost most of their ability to carry out a certain detoxification process. Continue to assume it is typical for a small percentage of liver cells to lose this ability in individuals in the reference class. In this case, I believe there would be little attraction to concluding that the liver or the cells are dysfunctioning, and scientists may disagree about how to classify the individual cells. In the end, I believe, this case would not be helpful for testing the FNC theory, due to uncertainty about whether the individual cells are dysfunctioning or not.

Finally, imagine that these cells' decreased ability is *unusual*: assume that the loss of ability occurs in only 2% of individuals in the reference class, perhaps due to some virus. In this case more than the last, I believe, there is some attraction to thinking that the individual cells are dysfunctional, even though the overall liver is not, and the FNC should be expected to come to the same conclusion. Note that even if it did not, this would form a relatively weak objection to the FNC since it's based on such a specific and unusual case, where our intuitions about the "right" answer are somewhat less reliable.[18]

But, on closer examination, the FNC gets the right answer. The cells are functioning at a level that is unusual in the reference class, so they satisfy the "frequency" part of the FNC's requirements. And, I would argue, the loss has a significant negative consequence as well according to the specifications presented above for how such consequences should be evaluated (Section 5.2). The negative consequence is the cells' inability

18. Here I am following the typical procedure in "conceptual analysis," checking proposed theories by their ability to conform to intuitions about the "correct" interpretation of a case. I have significant qualms about this general project, as do others, especially involving imaginary cases where "intuitions" are questionable (cf. Ramsey 1992; Schwartz 2004, 2007). See further discussion in Section 7 below.

https://doi.org/10.1086/521970 Published online by Cambridge University Press

to do something that has been standard for liver cells to do in individuals in this reference class for thousands of years. Just as in human blindness the eyes' loss of ability to see counts as a negative consequence whether or not this loss has an effect on survival and reproduction, so these liver cells' loss of ability to carry out the usual detoxification should count as a negative consequence even though there is no effect on the liver overall. Similar considerations apply to the case of the few dysfunctional sperm that Wakefield discusses.

In fact, cases like this can be slightly tweaked to form a challenge to the Frequency approach, since they regenerate the problem of healthy populations. Imagine that all liver cells in the reference class carry out the detoxification process at 80% of capacity or above; i.e., none are dysfunctioning. According to the Frequency approach, however, the bottom 2% or so (assuming normal distribution) would count as dysfunctional. The FNC account does not necessitate this outcome, since it allows for the relevance of a judgment that carrying out the detoxification process at 80% of capacity does not represent a significant decrement. Admittedly, there is some ambiguity how to define "significant" in these settings, but this may be unavoidable.

7. Conclusion. The FNC account, like the Frequency approach before it, does not provide an algorithm for drawing a specific line between low-normal function and dysfunction. There will always be some arbitrariness. But, the FNC account explains how to combine two factors—frequency and negative consequences, as specified above—to draw an acceptable line in many cases. This approach is superior to the Frequency approach, I believe, for the reasons described above. Which approach is chosen—one of these, or some other—is less important than recognizing and confronting the line-drawing problem. Failing to do so is a serious lacuna for any definition of function or dysfunction.

Philosophical analyses of concepts such as "dysfunction," "function," and "disease" are often framed as attempts at *conceptual analysis*, with the aim of uncovering the true meaning of a term (Boorse 1976) or the "criteria of application that people generally have in mind" (Neander 1991, 171). But these ways of defining the project rest on problematic assumptions about the nature of meaning (Millikan 1989) or the sort of criteria that speakers generally apply (Ramsey 1992). There is little question that in attempting to clarify and define concepts, we change them (Schwartz 2004, 2007). In this spirit, I believe that the FNC account should be evaluated as a possible approach for addressing the line-drawing problem in the future, rather than as describing how people have addressed

https://doi.org/10.1086/521970 Published online by Cambridge University Press

the line-drawing problem in the past.[19] (That said, I believe the FNC account does capture elements that are relevant to many experts' thoughts in this area.) The approach's availability for use in the future, and its success at matching current judgments in many cases, should help ease worries that drawing a line between functioning and dysfunctioning inevitably relies on problematic teleological assumptions or value judgments.

REFERENCES

Allen, Collin, and Marc Bekoff (1995), "Biological Function, Adaptation, and Natural Design", *Philosophy of Science* 62: 609–622.
Boorse, Christopher (1976), "Wright on Functions", *Philosophical Review* 85: 70–86.
——— (1977), "Health as a Theoretical Concept", *Philosophy of Science* 44: 542–573.
——— (1987), "Concepts of Health", in D. Van De Veer and T. Regan (eds.), *Health Care Ethics: An Introduction*. Philadelphia: Temple University Press, 359–393.
——— (1997), "A Rebuttal on Health", in J. M. Humber and R. F. Almeder (eds.), *What Is Disease?* Totowa, NJ: Humana Press, 1–134.
——— (2002), "A Rebuttal on Functions", in A. Ariew, R. Cummins, and M. Perlman (eds.), *Functions: New Readings in the Philosophy of Psychology and Biology*. Oxford: Oxford University Press, 63–112.
Bosch, J. L. H., et al. (1995), "Natural History of Benign Prostatic Hyperplasia: Appropriate Case Definition and Estimation of Its Prevalence in the Community", *Urology* 46 (Supplement 3A): 34–40.
Buchanan, Allen, et al. (2000), *From Chance to Choice: Genetics and Justice*. Cambridge: Cambridge University Press.
Caplan, Arthur (2004), "The 'Unnaturalness' of Aging—Give Me Reason to Live!", in A. L. Caplan and J. J. McCartney (eds.), *Health, Disease, and Illness: Concepts in Medicine*. Washington, DC: Georgetown University Press, 117–127.
Culver, Charles, and Bernard Gert (1982), *Philosophy in Medicine: Conceptual and Ethical Issues in Medicine and Psychiatry*. New York: Oxford University Press.
Cummins, Robert (1975), "Functional Analysis", *Journal of Philosophy* 72: 741–764.
Daniels, Norman (1985), *Just Health Care*. Cambridge: Cambridge University Press.
Dennett, Daniel C. (1995), *Darwin's Dangerous Idea*. New York: Touchstone, Simon and Schuster.
Gert, Bernard, Charles M. Culver, and K. Danner Clouser (1997), *Bioethics: A Return to Fundamentals*. New York: Oxford University Press.
Lewens, Tim (2004), *Organisms and Artifacts: Design in Nature and Elsewhere*. Cambridge, MA: MIT Press.
Lewontin, Richard (1983), "Gene, Organism and Environment", in D. S. Bendall (ed.), *Evolution from Molecules to Men*. Cambridge: Cambridge University Press, 273–285.
Lilienfeld, Scott, and Lori Marino (1995), "Mental Disorder as a Roschian Concept: A Critique of Wakefield's 'Harmful Dysfunction' Analysis", *Journal of Abnormal Psychology* 104: 411–420.
Matthen, Mohan, (1997), "Teleology and the Product Analogy", *Australasian Journal of Philosophy* 75: 21–37.
Millikan, Ruth Garrett (1989), "In Defense of Proper Functions", *Philosophy of Science* 56: 288–302.

19. I have called such a project *philosophical explication* and recommended it as an alternative to conceptual analysis in the discussion of definitions of "function" and "disease" (Schwartz 2004, 2007).

https://doi.org/10.1086/521970 Published online by Cambridge University Press

Neander, Karen (1991), "Functions as Selected Effects: The Conceptual Analyst's Defense", *Philosophy of Science* 58: 168–184.
——— (1995), "Misrepresenting and Malfunctioning", *Philosophical Studies* 79: 109–141.
Polvikoski, T., et al. (2001), "Prevalence of Alzheimer's Disease in Very Elderly People: A Prospective Neuropathological Study", *Neurology* 56: 1690–1696.
Ramsey, William (1992), "Prototypes and Conceptual Analysis", *Topoi* 11: 59–70.
Rettenmaier, J. L., et al. (2002), "Prevalence of Canine Hip Dysplasia in a Veterinary Teaching Hospital Population", *Veterinary Radiology and Ultrasound* 43: 313–318.
Schwartz, Peter H. (2004), "Moving beyond Conceptual Analysis in the Function Debate", *Monist* 87: 136–153.
——— (2007), "Decision and Discovery in Defining 'Disease'", in H. Kincaid and J. McKitrick (eds.), *Establishing Medical Reality: Essays in the Metaphysics and Epistemology of Biomedical Science*. Dordrecht: Springer, 47–64.
Wakefield, Jerome C. (1992a), "The Concept of Mental Disorder: On the Boundary between Biological Facts and Social Values", *American Psychologist* 47: 373–388.
——— (1992b), "Disorder as Harmful Dysfunction: A Conceptual Critique of DSM-III-R's Definition of Mental Disorder", *Psychological Review* 99: 232–247.
——— (1999a), "Evolutionary versus Prototype Analyses of the Concept of Disorder", *Journal of Abnormal Psychology* 108: 374–399.
——— (1999b), "Mental Disorder as a Black Box Essentialist Concept", *Journal of Abnormal Psychology* 108: 465–472.

https://doi.org/10.1086/521970 Published online by Cambridge University Press