'Rethinking "Disease": a fresh diagnosis and a new philosophical treatment'

Russell Powell,¹ Eric Scarffe

Department of Philosophy,
Boston University, Boston,
Massachusetts, USA

Correspondence to

Dr Russell Powell, Department
of Philosophy, Boston University,
Boston, MA02215, USA;
powell@bu.edu

RP and ES contributed equally.

Received 15 March 2019

Accepted 20 March 2019

Published Online First

2 July 2019

ABSTRACT

Despite several decades of debate, the concept of disease remains hotly contested. The debate is typically cast as one between naturalism and normativism, with a hybrid view that combines elements of each staked out in between. In light of a number of widely discussed problems with existing accounts, some theorists argue that the concept of disease is beyond repair and thus recommend eliminating it in a wide range of practical medical contexts. Any attempt to reframe the 'disease' discussion should answer the more basic sceptical challenge, and should include a meta-methodological critique guided by our pragmatic expectations of what the disease concept ought to do given that medical diagnosis is woven into a complex network of healthcare institutions. In this paper, we attempt such a reframing, arguing that while prevailing accounts do not suffer from the particular defects that prominent critics have identified, they do suffer from other deficits—and this leads us to propose an amended hybrid view that places objectivist approaches to disease on stronger theoretical footing, and satisfies the institutional-ethical *desiderata* of a concept of disease in human medicine. Nevertheless, we do not advocate a procrustean approach to 'disease'. Instead, we recommend disease concept pluralism between medical and biological sciences to allow the concept to serve the different epistemic and institutional goals of these respective disciplines.

INTRODUCTION

'Disease' is one of the most foundational concepts in medicine, and yet it is also one of the most intractable. Debates over the concept of disease are no closer to a resolution than when they began several decades ago. The dialectic is typically framed as a contrast between naturalism and normativism, with a hybrid view that combines elements of each staked out in between. There is still no consensus among medical theorists, ethicists or practitioners as to which of the three dominant accounts of disease on offer in the literature is preferable and why, nor is there broad agreement or even clarity on what the *desiderata* for a concept of disease ought to be. Given the well-known problems with existing accounts, one could be forgiven for thinking that the concept of disease is beyond repair and thus for recommending that it be eliminated in medical contexts.¹

Because discussions of 'disease' straddle biology, medicine, morality and culture, disease theorists have angled from many different disciplines in the service of diverse programmatic goals. This situation is conducive to talking at cross purposes and has caused some recent philosophical analyses of 'disease' to become rather insular,² quarantined

from wider social ramifications that one might reasonably insist should not be bracketed or ignored. Any attempt to reframe the debate must include a meta-methodological critique guided by expectations of what the disease concept ought to do, and how medicine as a discipline ought to be conceived, given that medical diagnosis is woven into a complex network of healthcare institutions. In this paper, we attempt such a reframing. Although critics are right to point out that prevailing views are in need of refinement, we argue that prominent critiques have misdiagnosed the most serious problems that plague existing accounts. It is in part due to this philosophical misdiagnosis that the debate has become so intransigent and compartmentalised. In what follows, we discuss the chief criticisms of each theory of disease and explain why these critiques gloss over the most powerful objections. We then go on to offer a new twist on the hybrid account that avoids these objections and satisfies the institutional-ethical *desiderata* for a concept of disease in human medicine. Finally, because the disease concept is both salvageable and ethically defensible, we conclude that an 'eliminativist' position is both unwarranted and undesirable. Nevertheless, we do not advocate a procrustean approach to 'disease'. Instead, we recommend disease concept pluralism between medical and biological sciences in order to allow the concept to serve the different goals of these respective disciplines.

1. IS NATURALISM FREE OF MORAL VALUES?

Let us begin with disease naturalism. Common to all naturalistic accounts is that they propose definitional criteria that are purported to be purely 'descriptive'. What they purport to describe are underlying biological matters of fact, not evaluative judgements. The relevant biological matters of fact are typically glossed in terms of some biotheoretic understanding of 'dysfunction'. Whether naturalistic accounts are vulnerable to a variety of objections depends, in part, on how biofunction is conceived. There are several prevailing understandings of biofunction that might be imported into the medical context, and we will say more about them in a moment. The takeaway for now is that according to naturalistic accounts, whether a trait is dysfunctional is determined by facts about the biological world, not the interests of researchers, the goals of policy makers or the desires of patients. Dysfunction can be disvalued and mitigated, and it often is; function can be valued and promoted, and it often is; but these value and policy judgements do not enter into the descriptive analysis of function, and thus they do not figure in the naturalistic

- <http://dx.doi.org/10.1136/medethics-2019-105555>
- <http://dx.doi.org/10.1136/medethics-2019-105578>
- <http://dx.doi.org/10.1136/medethics-2019-105591>
- <http://dx.doi.org/10.1136/medethics-2019-105663>
- <http://dx.doi.org/10.1136/medethics-2019-105745>

© Author(s) (or their employer(s)) 2019. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Powell R, Scarffe E. *J Med Ethics* 2019;**45**:579–588.

account of disease. So it is that naturalistic theories appear, at least on their face, to be moral value-free.

Nevertheless, an influential line of criticism holds that although naturalistic accounts claim to offer a value-free definition of disease grounded in 'objective' biological facts about function, they wind up importing values after all. Of course, values of the epistemic variety will always underlie the specification of concepts and adjudication of theories in the sciences.³ The charge against naturalistic theories of disease, therefore, is that they unwittingly incorporate *moral* values, or evaluative judgements that advert to well-being, flourishing, opportunity and the like. The assertion is that in concept and/or in practice, disease classifications are shaped, often insidiously, by covert value judgements of the robustly normative kind.

For example, for much of the 20th century, homosexuality was considered a disease by the American Psychiatric Association (APA) and WHO. The fact that homosexuality is no longer considered a disease appears to reflect a shift in social values, not a new understanding of biofunctional facts about the trait in question.^{4–7} Although some highly speculative hypotheses have been advanced that propose homosexuality has adaptive origins, this was clearly not the underlying motivation for the declassification of homosexuality. The reason why the APA and WHO no longer classify homosexuality as pathological is not because adaptationist hypotheses have prevailed, but rather because of the triumph of new patterns of social evaluation. Some theorists even argue that how an account classifies homosexuality should serve as a litmus test for its philosophical viability.⁷

This critique gains some general traction when one realises that the homosexuality case is far from an isolated incident: there are in fact numerous examples where disease classification has clearly tracked changes in social values, as is the case with masturbation⁸ in the 18th and 19th centuries, and with various psychiatric diagnoses in the 20th century. Indeed, the most recent iteration of the *Diagnostic and Statistical Manual of Mental Disorders* (released in 2013) classifies the fear of public speaking, nail biting and depression immediately following the death of a loved one as medically treatable pathologies, even though these conditions do not intuitively constitute diseases. This suggests that properly moral values, not biological facts, are driving the 'pathologisation' of normal variation, especially where this classification has implications for how medical resources are to be allocated (see §4).

This critique does little to undercut the naturalist, however, who can simply reply that these cherry-picked (although admittedly numerous) cases merely reflect occasional political influence in what is otherwise a value-free classification scheme. In other words, the naturalist can readily reply that these disputed cases are not genuine philosophical 'counterexamples', which show that the necessary and sufficient conditions cited by naturalist accounts fail. At best, the response contends, they show that *in practice* disease classification has sometimes been shaped by moral value judgements that have resulted in erroneous classifications. One might reply that normative considerations factor into the argument that 'disease' *ought* to be conceived in moral value-neutral terms, and hence that this position is itself a value judgement. Whether this value judgement is a properly moral one is unclear, however, as this will depend on the nature of the reasons that are offered in its favour. This is our first indication that the intractability of the 'disease' dispute may stem from a fundamental lack of clarity about, or tacit disagreement over, the relevant desiderata for a medical conception of disease.

In any case, given the responses available to the naturalist, a more effective line of attack would be to show that disease

naturalism is incoherent or fails on its own terms because it is not grounded in biology alone. If persuasive, this argument could be marshalled into a forceful criticism of hybrid theories as well insofar as they incorporate a 'natural' bio-dysfunction component. The charge that naturalist accounts of disease are infused with moral value, notwithstanding their protestations to the contrary, is a long-standing one. But this criticism takes on greater urgency in recent years as some prominent philosophers of science have charged disease naturalism with presupposing value-laden goals and archetypes—the ultimate sin of post-Darwinian biology. These criticisms have typically been aimed at Boorse's 'biostatistical' view,^{9–10} which is by far the most influential naturalistic account of disease. We will briefly summarise these criticisms and explain why they do not, in fact, make contact with Boorse's theory—and that, even if they do, they fail to generalise to other accounts that repose on a proper evolutionary understanding of function.

Boorse's theory attempts to connect the diverse range of phenomena that are commonly taken to be diseases, from infections and cancers to limb paralyses, with a common theoretical thread. He argued that despite the diversity of their underlying causes, all of these conditions implicate biological dysfunction. He thus introduces the notion of 'normal function', which he cashes out in terms of a statistically typical contribution of some part or process in individuals of a given reference class to survival and reproduction. When traits depart from normal functional efficiency by some stipulated degree, they are deemed dysfunctional—and hence diseased—against the idealised standard of what he calls 'normal species design'. Of course, if the stipulated departure from normal functional efficiency is drawn on the basis of properly moral values (such as well-being), then the account will cease to be 'naturalistic' in its intended sense. However, grey areas pervade the living world, and ontological boundaries can nonetheless be drawn without moral values figuring into the demarcation.

A more potent criticism is that there are no naturalistic resources for Boorse to draw on that can tell us what the 'normal' traits of a species might be. This is because species are historical entities with their parts (populations, organisms) united by common ancestry—not natural kind classes that organisms belong to because they possess certain essential characteristics.¹ Moreover, trait development is the result of 'norms of reaction', or interactions between genotypes and complex non-genetic environments, as well as epistatic interactions between genes—not the operation of internal, unseen essences.^{11–13} Although we can identify different statistical patterns of trait development—that such and such genetic backgrounds, interacting with such and such non-genetic environments, tend to results in such and such distributions of outcomes—we cannot say which of these patterns is 'normal' for a given species. Boorse's view thus appears to rest on some idealised standard or 'archetype' against which dysfunction is to be measured, and since there are no naturalistic resources he can draw on to set this idealised standard, moral values must be factoring into the archetype. The charge that Boorse's view presupposes a value-laden archetype is a common objection to disease naturalism, one that is commonly levelled by ethicists. Sisti *et al*,¹⁴ for instance, dismiss disease naturalism because it 'reflects the belief that there exists a Platonic world in which species, classes and kinds can be objectively distinguished by identifying their pure, essential properties'.

Boorse's unfortunate appeal to 'normal species design' does ring of archetypes. However, as we argue elsewhere,¹⁵ Boorse's theory is at odds with neither the received view that species are individuals (rather than natural kind classes) nor with the modern

biological rejection of Aristotelian essentialism. The regularities on which Boorse's notion of 'species design' depends amount to no more than statistical regularities produced and maintained by thoroughly Darwinian causes, such as common ancestry, stabilising selection and sexual recombination. Although Boorse's account of disease falters for other reasons (see below and §3), we do not think it can fairly be faulted for relying on a pre-Darwinian notion of species. Nor do we think, as some philosophers of science have argued,¹ that Boorse's view reposes on morally normative assumptions because it imputes of the goal of survival and reproduction to organisms. The mere fact that a trait is an adaptation does not imply that we as critical evaluators ought to value it. However, nowhere does Boorse claim that survival and reproduction is a goal of human beings *qua intentional, rational agents*. The goals that Boorse has in mind are biological, not psychological—and this explains why he is adamant that disease may or may not detract from human life in ways that warrant treatment, and why his account is touted as universal rather than limited to organisms that can value things. There may be *epistemic* values that figure in Boorse's choice of goals—such as the inferential or explanatory role they play in the relevant scientific enterprises—but there is nothing self-defeating or non-naturalistic about this.

Many of Boorse's problems stem from the fact that although the concepts he uses to delineate function gesture at evolutionary principles, they are only weakly evolutionary. As a result, his account is saddled with miscommunications and ontological difficulties that arise within a conceptual framework that in effect 'branched off' before the evolutionary concept biological function was rigorously developed in biology and philosophy of biology over the last four decades of research. Let us grant for the sake of argument that Boorse's account fails on naturalistic grounds. This does not condemn disease naturalism more broadly to the same fate, since the naturalist can avoid all of the difficulties mentioned above by adopting an explicitly evolutionary account of function: one that adverts to a history of selection for one or more effects.^{16–17} These selected effects are deemed evolutionary functions, and sufficient departures from these evolutionary functions constitute dysfunctions.

Because it is firmly entrenched in contemporary evolutionary theory, the selected effects account of function comes with a host of explanatory and operational advantages that make it better suited to support a naturalistic theory of disease. It is immune from the 'value-laden archetype' criticism (since it adverts neither to 'normal species design' nor to 'goals'), and it avoids a slew of pseudo-problems that arise from the failure of the Boorsean account to index function to evolutionary timescales. We discuss these advantages elsewhere and will not delve into them here.¹⁵ The upshot is that prevailing criticisms of disease naturalism are either addressed to straw men, or else succeed at dismantling one particularly dismantle-able view, but do little damage to evolutionarily accounts that are made of sturdier theory.

Substituting evolutionary function for Boorsean normal function may render 'disease' moral value-free, but it does not save disease naturalism. Although the selected effects account of function avoids many of the criticisms that (rightly or wrongly) beleaguer Boorse, it, too, fails to meet the *desiderata* for a concept of disease in human medicine. Naturalism's biggest selling point—its value-independence—is also its biggest flaw. We will argue that naturalistic accounts do not succeed for the very same reason that explicitly value-based theories do not succeed: namely, they both fail to capture the institutional and ethical dimensions of the disease concept.

2. Value-based accounts and their limitations

Disease normativism maintains that classifications of 'health' and 'disease' reflect value judgements about biomedical states of affairs and nothing more.^{18–21} Diseased biomedical states are simply those we disvalue or wish to avoid. Normativism is confronted with a series of well-known problems. First, it renders disease classification spatiotemporally contingent: different cultures will value diseases differently, and thus the same state can be a disease in one culture but not in another; and since patterns of evaluation change within societies over time, the same state can be a disease at one time in a culture and not at another time. The normativist is not troubled by these contingencies, however, since they maintain that diseases are mere social constructions, not natural kinds, and hence they are happy to jettison misguided notions of objectivity that have obscured the subjective, values-based nature of disease classification.

The normativist has also been challenged by various counterexamples which purport to show that disvaluing a biomedical condition is insufficient to classify that state as a disease (take, for instance, pattern baldness or diminutive male sex organs)—but like the naturalist, the normativist is unmoved by a few cherry-picked examples. As evidence for disease subjectivism, normativism points to politically and religiously motivated cases of disease classification, such as the pathologisation of political dissidence,^{22–23} the resistance to oppression, the desire of slaves to flee captivity ('drapetomania') and females having sex outside of marriage, to name a few. Indeed, some prominent sceptics of 'disease' have argued that mental illness classifications in the 20th century have been widely motivated by similarly manipulative sociopolitical goals.²⁴ Normativism's detractors, meanwhile, complain that a major problem with normativism is precisely that it does not have the resources to say that such morally problematic cases constitute erroneous applications or abuses of the disease concept, which many would take to be a problem with, not a strength of, the normativist position as it is canonically formulated. A more robustly normative normativist view does a better job handling these cases, and we discuss such views, and problems with them, later on in this paper (§4).

These problems with normativism as classically conceived have led some disease theorists to try to square the circle by combining the biological and social elements of disease. Incorporating a biofunctional element into the concept of disease restricts its range of application, while including a values element accords with how the disease concept has been and continues to be used—perhaps ineluctably—in medical practice. This merger is achieved by hybrid approaches, of which Wakefield's²⁵ version is the most influential. Wakefield proposes that disease be equated with 'harmful dysfunction', where 'harmful' is an evaluative term determined by social values, and 'dysfunction' is a scientific term referring to the failure of a mechanism to perform its evolutionary function. The thought is that the conjunction guards against abuses by tying the subjective normative component to an objective biological one. We can readily see that 'traits' such as political dissidence, drapetomania and masturbation do not satisfy the biological dysfunction component regardless of whether a particular society deems them to be harmful.

How have critics responded to the hybrid view? Some have sought to undermine the biological component using the same set of criticisms that are thought to defeat naturalism.¹ However, we have already seen that Boorse's normal function model does not suffer from the problems these critics have attributed to it, and that even if it did, a selected effects account could replace it at

little cost and much gain. Still other critics have sought to undermine the normative component of the hybrid account, arguing, for instance, that normative and hybrid accounts fail because physicians can (and often do) find their patient to be unhealthy even when their patient does not subjectively disvalue their medical state.²⁶ Although this objection may land on subjectivist views that deem individualised evaluation a necessary condition for disease,²⁷ it misses the mark more generally because the evaluation incorporated into most value-based accounts of disease is societal (or, as we shall argue in §4, idealised), and thus abstracts away from interpersonal variation in subjective evaluation. Naturalists, of course, will remain unhappy with contingent social facts determining—even in part—whether a biomedical condition is a disease.

Although the hybrid view's criticisms of normativism are well-received, they fail to identify normativism's key flaw: namely, that rather than being *too* value-laden, existing value-based accounts are *not normative enough*. Normativist and hybrid views are indexed to how biomedical states are valued, but they make no claim about whether these states are *properly valued as such*. Recall the homosexuality case: naturalists run into trouble here because homosexuality looks straightforwardly dysfunctional from the standpoint of biological fitness. Naturalists will be keen to stress that by saying 'homosexuality is a disease', they are not implying that it should be disvalued or that the medical community should attempt to treat it. But not surprisingly, this is of little comfort to many theorists and activists. On normativist and hybrid views, homosexuality was correctly de-classified because it was no longer disvalued by the relevant populations. Although normativist and hybrid views pass the homosexuality litmus test, they do so for the wrong reason. Homosexuality should be declassified not because a certain level of consensus has been reached about how it ought to be valued, but rather because disvaluing homosexuality is not rationally justified and causes objective harm and injustice. In other words, truly normative ethics (whether centred around flourishing, well-being, or a broader range of values) is needed to rein in weakly normative theories of disease.

The real problem with hybrid accounts, then, is that, like normativist theories, they are not 'normative' in the true sense of the term. Whether a biomedical state constitutes a disease on such views depends on bare facts about evolution and social evaluation, not on any properly normative considerations. This problem is reflected in how the hybrid account handles socially conservative societies that demonise homosexuality. Wakefield insists that such societies are misapplying the concept of disease because of the contingent (and in our view, clearly disputable) fact that there is no underlying bio-dysfunction after all.²⁸ If this is right, then whether homosexuality should be classified as a disease in homophobic cultures turns wholly on our evolutionary verdict regarding the underlying proximate mechanisms. Thus, although weakly normative hybrid accounts can make sense of misapplications that do not implicate biological dysfunctions (such as masturbation or drapetomania), they do not have the resources to recognise as misapplications biological dysfunctions that are wrongfully or perniciously disvalued. Other articulations of the hybrid view suffer from a similar defect.²⁹

Thus, in a sense, all existing views of disease are purely descriptive: some describe biological facts, others describe social facts and still others describe a conjunction of biological and social facts. This deficit is an artefact of the methodology that has typically been employed in the 'disease' literature. Naturalist, normativist and hybrid accounts have all been engaged in the traditional philosophical project of conceptual analysis: they

are asking what some community of language users—be they physicians, pathologists or patients—means by the term 'disease' when they use it. The method of reflective equilibrium is then used to reconcile philosophical accounts against intuitions in specific cases. In fact, what is needed is a theory of disease that takes normativity seriously, and takes *institutional context* seriously. We will use the remaining space to outline such a theory.

3. The normativity thickens: a new hybrid view

Our proposed definition of disease is as follows: a biomedical state is a disease only if it implicates a biological dysfunction that is, or would be, properly disvalued. This might appear to be a minor modification to existing hybrid theories, but the injection of thick normativity confers major advantages onto the hybrid view.

First, a thick moral evaluation avoids the spatiotemporal contingency problems that afflict normativist and weakly normative hybrid views, since objective normative standards are not affected by social patterns of evaluation that change over space and time. Of course, our *views* regarding what should be disvalued may change; but whether a biomedical state is a disease does not change in lock step with social shifts in evaluation. This metaphysical stability is due to two objective components: biological dysfunction and rational moral justification. We showed earlier that bio-dysfunction can be delineated without recourse to *moral* values, particularly if a selected-effects account is substituted for Boorsean normal function (see §1). The view that moral norms can be objectively justified is held by most contemporary thinkers in moral and political philosophy who are not moral sceptics, including moral rationalists and prominent feminist philosophers of science,³⁰ cultural moral relativists and non-cognitivists.³¹ Furthermore, claiming that moral norms can be objectively justified need not commit oneself to a strong species of moral realism, or to the notion that morality is mind-independent.^{32 33} Given widespread philosophical and bioethical agreement on these points, it is all the more surprising that an unapologetically objective, value-based account of disease has rarely been defended.

What uncontroversial things can we say about the 'rational justification' element, given that the methods of moral justification are contested? At the very least, we can say that reasons independent of popularity and tradition must be adduced to justify the evaluation. We can say that these reasons must not be arbitrary or bigoted, that they should not repose on false empirical claims or inscrutable religious beliefs and that they must be subject to critical scrutiny and revision. Many cases of pernicious disvaluing are motivated by faulty social medical epistemologies that involve motivated reasoning and rest on false empirical beliefs. For instance, some historically argued that masturbation was bad because it brought about other known maladies (such as gonorrhoea), and religious mores may have made these false beliefs more salient or resistant to falsification.³⁴ Likewise, disgust reactions continue to fuel derogatory moral attitudes towards homosexuals and other non-conforming sexual practices and identities, as does religious dogma that is impervious to critical examination.

Relatedly, we can also say that the supposed 'naturalness' of a biomedical state cannot be among the acceptable reasons offered in support of an evaluation. There is, of course, the standard spate of problems delineating natural from unnatural biological outcomes (see §1), as well as the naturalistic fallacy which derails any attempt to move from the property of unnaturalness to the property of badness without an independent line of argument to fill this gap. Quite apart from these obvious problems with

naturalness, however, if we did not exclude such reasons, then *all* biological dysfunctions would be properly disvalued, insofar as all biological dysfunctions are unnatural. We thus agree with the disease naturalist that whether a given bio-dysfunction is a proper target of medical treatment can only be determined in reference to moral values that are independent of facts about dysfunction and unnaturalness themselves. Even if the existence of bio-dysfunction creates a *de facto* presumption of proper disvaluing, this presumption may be over-ridden if ethical arguments make a sufficiently strong case for the contrary.

Of course, a thickly normative hybrid view will not appeal to those who are incorrigibly sceptical about the possibility of moral objectivity in any form. But it should be attractive to anyone who thinks that moral justification is a reason-giving process, and that justificatory appeals to authority, popularity, theological premises or false empirical claims will not suffice, especially where these appeals are intended to inform law and policy.

What else can be said about rational moral justification? Another thing we can say is that moral consensus *per se* is not a good indicator of sound justification. We agree with Sisti *et al*¹⁴ insofar as they hold that values do—and should—shape diagnoses of disease, and that acknowledging this value-ladenness does not resign us to subjectivity or arbitrariness in our diagnoses. But while Sisti *et al* are right that certain core values, such as self-determination and human dignity, currently enjoy widespread recognition (a few loud dissenters notwithstanding^{35 36}), the fact that this consensus exists is not what justifies these values or secures their objectivity. Just as biological dysfunctions remain dysfunctions even if they become statistically normal in a population, the rational justification of a particular moral evaluation is unaffected by popular adherence to, or divergence from, that evaluation. More important than the popularity of a norm, from a moral epistemic perspective, is the nature of the processes through which the norm developed and proliferated: did that process exclude the perspectives of moral stakeholders, and did it take into account the best empirical information by identifying genuine sources of scientific expertise?

This view of rational moral justification is consistent with the notion that the content of rationally defensible moral norms is to some degree contingent on background institutions. The validity of moral norms cannot be assessed in an institutional vacuum—this is as much true for moral evaluations about biomedical states as it is for norms regarding the use of force at individual and state levels. For example, the appropriate content for a norm of self-defence may depend on the institutional recourses that are available for the prevention of aggression.^{37 38} Likewise, a given biological dysfunction (eg, dyslexia) may be properly disvalued in certain institutional contexts (eg, one in which literacy is crucial to partaking in the dominant cooperative scheme), but not in others. The fact that institutional context is relevant to proper moral evaluation does not undermine the possibility of rational moral justification—it merely shows that the latter must take institutional considerations into account in adjudicating moral claims. However, as we shall see, taking socioinstitutional context into account leads to problems when the context is itself unjust.

Somewhat more tentatively, we might say that moral evaluation should be based on considerations that are limited to the flourishing of the individual with the biological dysfunction in question. In other words, the evaluation should not take into account indirect costs for society (be these social, economic, aesthetic or what have you). One reason to think the justification ought to be limited in this way is that the central goal of

contemporary human medicine is not to serve the non-health goals of society, but to enable patients to make informed, autonomous decisions about their health and course of treatment in a way that is consistent with their own values. This is not to essentialise the purpose of medicine, but simply to say that our best current ethical thinking about medicine prioritises patient self-determination over collective social interest and even over the medical duty of healing. Allowing indirect costs to creep into the evaluation creates a susceptibility to prejudice, unfairness and tyranny of the collective good.³⁹

In some cases, however, individual flourishing or well-being may be shaped by aberrant subjective preferences that pose a significant risk of harm to others, such as in cases of paedophilia and psychopathy.⁴⁰ In such instances, preventing objective harm to others could justify a negative moral evaluation that results in classifying these conditions as diseases (insofar as they implicate bio-dysfunctions), or as non-diseased states that otherwise warrant biomedical intervention. Nevertheless, the problem of aberrant subjective preferences underscores the bioethical limitations of a parochial focus on the flourishing of individual patients.

A harder question, flagged above, is how to treat cases in which the conditions for individual flourishing are themselves shaped by unjust (eg, bigoted) social norms. In such scenarios, labelling certain conditions ‘diseases’ could reinforce existing injustice. For instance, some conservative groups in the USA have argued that being gay is associated with an increased risk of mental illness, drug abuse and relationship instability, and ought to be considered a treatable disease on that basis. But the reason, of course, that these causal associations obtain (if they do) is that there remains a pervasive social stigma associated with homosexuality and there is far less social support for lesbian, gay, bisexual and transgender populations than there are for heteronormative communities. This is even more the case in virulently homophobic and transphobic societies. Thus, if we look solely at the individual prospects for flourishing without considering the unjust social environment in which this flourishing plays out, then we might say that being gay is properly disvalued and that gay conversion therapies could be warranted, whether on a disease-based or ‘welfarist’ (§4) view of the justification for medical intervention. To avoid exacerbating or endorsing unjust norms and institutions, therefore, the moral evaluation of biomedical states should take into account the legitimacy of the social environment in which individual flourishing takes place. This proposition will seem even more plausible once it becomes clear that ‘disease’ is an inherently moral concept that is motivated by considerations of justice, as we will argue in the next section.

Having said this, in societies or subpopulations in which the conditions for flourishing and opportunity for those with a certain gender identity or sexual orientation are severely restricted—for example, where such individuals are reliably subject to extreme physical violence and social punishment that would prevent them from living a good life—it is not unreasonable to hold that these biomedical states are rightfully disvalued (in a very narrow, contextual sense) by individuals who have them and thus could be a proper target of medical intervention. In such cases, there is a tension between acting in furtherance of individual flourishing, on the one hand, and combating social injustice, on the other. Precisely how this tension should be resolved in medicine is unclear, and reasonable people may differ on this matter. We maintain that in these tragic scenarios, medical intervention, if it is to be carried out at all, should not be carried out under the banner of treating a ‘disease’. The upshot is that individual

well-being/opportunity cannot exhaust the moral analysis—the broader socioinstitutional context must also be considered.

A second advantage of a thickly normative hybrid account is that it can make sense of misapplications and pernicious abuses of the disease concept in the past, while recognising that there has been progress in our understanding of which biomedical states constitute diseases as new normative and biological information has come in. For instance, some misapplications of the disease concept have stemmed from ‘imperialistic’ conceptions of human good that were tinged with racial, ethnic and gendered subtexts, as well as from a general failure to appreciate the lived experiences of disabling conditions. A growing appreciation for value pluralism—a recognition of the diversity of reasonable perspectives on the good life—has reshaped our understanding of which biomedical states are objectively disvaluable. Pluralistic approaches to medical ethics have emerged in recent decades, stressing mutual respect for divergent and often conflicting values in medical decision-making and policy.⁴¹ Our failure to recognise the reasonableness of diverse values, or of the weighting that is accorded to different values, has historically resulted in problematic moral evaluations of biomedical states and ethical abuses that have flowed therefrom. At the same time, there has been much progress on the mechanistic front, with the discovery that certain ‘cognitive’ conditions—such as addiction, mood, attention and post-traumatic stress disorders (box 1)—have neurodysfunctional causes, reducing their social stigma and bringing them more firmly within the ambit of medicine.

A third advantage of our proposed account is that it interacts well with disability rights critiques that see the disease concept literature (especially as it is framed in the philosophy of science) as glossing over the first-person phenomenology of patients.^{42–43} Taking into account these lived experiences might prompt us to rethink how certain conditions are (dis)valued. Conditions from deafness to autism have been reclaimed by some patients and theorists as valuable expressions of human diversity, rather than as diseases to be mitigated or eliminated through the powers of modern medicine. Although many disability rights critiques have targeted the ‘normal function’ component of naturalistic disease models,⁴⁴ given the resilience of evolutionary (selected-effects) function to these objections (see §1), a more successful point of critique is at the locus of moral evaluation itself. Whatever one thinks about the ‘reclamation’ of various disease states (for a critique of such arguments, see McMahan⁴⁵), the resources for making these arguments, and for countering them, are contained in the thickly normative component of the proposed view.

This brings us to the biggest advantage of the account we are proposing: namely, it is well-suited to the role that the disease concept plays in developed healthcare institutions. Because it takes institutions seriously, our proposed view does not attempt to sever disease classifications from their social ramifications, as naturalistic theories are wont to do. In response to various objections from clinicians, Boorse has cabined his theory to the point that it is now limited to describing how pathologists are using the term, effectively insulating his view from social-institutional critique. Even this parochial endeavour may not succeed, for as Wakefield⁴⁶ points out, pathologists appear to distinguish ‘pathological’ states from ‘diseases’ proper. Existing hybrid views speak to a broader community of language users and do a more admirable job of handling the clinical aspects, but they still do not quite capture the thickly normative dimensions of the disease concept as it is deployed in contemporary healthcare institutions. The methodological approach we employ here is not conceptual analysis, but rather *conceptual role* analysis—an inquiry shaped less by linguistic usage, and more by the role that

Box 1 Case study: post-traumatic stress disorder (PTSD)

Disvalue

Most salient somatic diseases—such as infectious, cardiovascular, oncological, immunological and neurodegenerative diseases, as well as various organ system malfunctions—have straightforwardly negative implications for human well-being. Although there are some morally contested cases in the category of mental illness, many psychiatric disorders caused by biologically defined dysfunctions will have psychological and social consequences that are clearly disvaluable on any plausible account of human well-being or flourishing. For instance, people who suffer from PTSD tend to experience intrusive memories of traumatic events, insomnia, nightmares, hypervigilance and hypersensitivity to ambient stimuli, protracted episodes of emotional dissociation and social detachment, emotional numbing and an inability to form stable attachments and to sustain intimate relationships.⁴⁵ One particularly morally relevant feature of PTSD, especially when the disorder is caused by chronic sexual abuse or childhood trauma, is the shattering of the victim’s personal sense of agency as a result of their violent domination by a more powerful individual. As the psychological inertia of past trauma continues to shape one’s life, one’s sense of agency is eroded yet further, and the trauma continues to be (re)lived. The long hand of trauma can also have significant deleterious impacts on physical health over the course of a lifetime, both directly through chronic stress response and indirectly by precipitating unhealthy behaviours. These are all significant impediments to the exercise of individual autonomy and the ability to flourish in our contemporary social world. One might argue (implausibly) that the exercise of autonomy or human flourishing are *not* properly normative considerations; or, one might argue (also implausibly) that although these are properly normative considerations, PTSD does not have a significant enough impact on them to rise to the level of being *properly disvalued*. Notice, however, that any such objections would turn on a reason-giving process of moral justification.

Dysfunction

Although the precise neural mechanisms underlying PTSD are poorly understood, researchers have identified several neurobiological deficits that are strongly indicative of emotional dysregulation. Traumatic stressors result in a number of cognitive-affective dysregulations, some of which are common to anxiety disorders while others look to be specific to PTSD. These include exaggerated amygdala and insula activity which modulate fear and startle responses; as well as hypoactivation in brain regions associated with the processing of negative emotions (including the ventromedial prefrontal cortex, rostral and dorsal anterior cingulate cortex, thalamus and the dorsal amygdala), which results in the autonomic blunting associated with episodes of numbing and dissociation that are characteristic of PTSD.⁵¹ Thus, PTSD appears to satisfy the moral and biological criteria of our proposed thickly normative hybrid account of disease.

the concept plays, and should play, in the institutional environments in which it is embedded. Because the technical disease concept is deployed well beyond scientific communities (let alone the more restricted realm of pathologists), it is critical that this wider institutional context be taken into account in determining how the disease concept ought to be specified and what desiderata it must meet.

4. AN INSTITUTIONALLY SENSITIVE CONCEPTION OF DISEASE

Eliminating either the bio-dysfunction component or the thick evaluation component of 'disease', or doing away with the concept of disease in its entirety, would have significant ethical costs that are generally underappreciated in the literature. These costs become clear once we recognise that disease diagnosis is not confined to an intellectual scientific community. It is a practice that is embedded in a diverse network of public and private institutions tasked with research, risk-pooling, regulation, allocation and delivery. Healthcare systems include biomedical research sectors, managed care organisations and provider networks, all of which are interacting with patients and one another in legal-regulatory regimes designed to coordinate the delivery of healthcare. The overarching goal of a healthcare system is to manage biomedical states of affairs in large, demographically diverse populations, and to do so in a manner that's consistent with basic ethical principles. In this complex institutional framework, the classification of certain biomedical states as diseases does not merely carry information about the right to some form of social biomedical support; it also helps to ensure that the allocation of medical resources is carried out efficiently and fairly at the population level by prioritising segments of the population who tend to be the worst off, biomedically speaking.

The disease concept provides a consistent and operational heuristic by which to prioritise treatment over enhancement. Buchanan *et al*⁴⁷ argue, persuasively in our view, that adhering heuristically (although not categorically) to a treatment-enhancement distinction enables us to prioritise biomedical conditions that tend, on average, to have the greatest impact on well-being and equal opportunity. Biomedical states vary in their impact on well-being and opportunity, and the same state is bound to vary in its impacts from individual to individual. Infertility, for example, may be a convenient form of birth control for some, but for others who want to have children it has significant disvalue. These impacts cannot be assessed on a case-by-case basis at the population level, given the obvious practical impediments to doing so. Anyone who has tried to contact public or private medical insurers to query a coverage decision is painfully aware of their very limited capacities for individual-level engagement. By abstracting away from interpersonal variation and the quirks of individual cases, the disease concept can serve as a population-level heuristic that tracks statistical, ethically relevant differences in classes of biomedical variation.

The argument for prioritising the treatment of disease over interventions that target normal, low-end variation goes something like this: if biological dysfunctions tend, on average, to interfere with well-being and opportunity more so than low-end variation that does not rise to the level of dysfunction, and if we have reason to prioritise the medically worst-off in our allocation of limited healthcare resources, then we have pragmatic reasons for adhering to a treatment-enhancement distinction that is drawn in part on the basis of biological dysfunction. If one rejects the notion that healthcare allocation ought to prioritise the worst off (medically speaking), then this removes one weighty reason to adhere to the treatment-enhancement distinction. However, many accounts of justice, from liberal egalitarian to utility maximisation models, do plausibly lead to this conclusion about prioritisation, and developed healthcare systems as a whole tend to adhere to this position. Although Buchanan *et al* rely on Boorse's 'normal function' model to draw the treatment-enhancement distinction, we believe that a thick value component is also crucial, since biological dysfunctions that

are not properly disvalued should not be subject to resource allocation.

This focus on treatment is not meant to preclude medical interventions that target non-disease states or the enhancement of normal function. Disease diagnosis is not the be-all and end-all of healthcare allocation decisions, since disease does not exhaust the range of biomedical states that interfere with well-being, flourishing or opportunity. Many conditions that do not implicate disease may nonetheless be afforded a high priority in medical institutional support, such as access to contraception and abortion services. And, to the extent they become safe and effective, some enhancements of lower-end function (such as cognitive enhancements) may have major impacts on well-being and opportunity that justify affording them a high priority in the allocation of biomedical resources. Including cognitive enhancements in social medical allocation may be especially important if society transforms in ways that cause lower-end variation of that trait to become, in effect, a disabling condition. Still, the disease concept does a reasonable job of statistically tracking biomedical conditions that tend to have an impact on these morally relevant outcomes, and thus it provides a non-arbitrary, consistent baseline for healthcare allocation decisions. It is thus no accident that all developed healthcare systems rely on the treatment-enhancement distinction in distributing limited medical resources, and any conception of disease should be well-suited to serve this institutional purpose.

Some authors dispute the ethical import of the treatment-enhancement distinction, arguing that allocations should be made by looking directly at the expected effects of medical interventions on well-being regardless of whether they target disease. As two recent proponents of this 'welfarist' view, Roache and Savulescu, contend, in considering whether to use medicine to treat a condition, 'whether that condition is characterised by biological pathology, or otherwise classifiable as a disease, (is) irrelevant'.⁴⁰ Instead, they maintain, we should focus solely on well-being and make healthcare prioritisation decisions in what amounts to an a-disease healthcare environment. Similarly, Harris and Chan⁴⁸ argue that decisions about the appropriateness of medical intervention should be based solely on risk-benefit calculation, not on whether there is an underlying disease to treat. To the credit of the welfarist view, jettisoning a bio-dysfunction component has the added benefit of side-stepping some of the difficulties that beleaguer all biofunctional accounts of disease. For instance, it avoids the critical unresolved problem of postreproductive diseases (or what Harris and Chan call the 'depredations of old age'), which arguably represent the greatest challenge to naturalistic theories of disease whether these accounts are indexed to evolutionary or normal function. The problem, simply put, is that many seemingly uncontroversial diseases associated with senescence likely result, at least in part, from the *lack of selective investment* in tissue maintenance and repair in postreproductive phases of the lifespan.⁴⁹ It is not immediately apparent, therefore, how these biomedical states could be categorised as diseases on a model of function that is indexed to selection, regardless of the historical timeframes involved. We will not attempt to solve this problem here; we raise it simply to show that the project of tying bio-dysfunction to disease remains incomplete even on our best understandings of biological function, and this is a point in favour of a functional welfarism.

We do not wish to make any pronouncements about the ultimate viability of the welfarist approach, which we take to be motivated by concerns that largely overlap with our own. Whether medicine and the broader healthcare apparatus can

function effectively if it is reconfigured to avoid the disease concept entirely is a question that we will revisit in the final section. But the fact remains that, as they are currently configured, all developed healthcare systems rely heavily on the treatment-enhancement distinction in determining medical necessity; and so, at least in the current institutional environment, any concept of disease should be well-tailored to this purpose.

Existing accounts of disease, as we have shown, are ill-suited to this task. Naturalistic accounts simply ignore the social and ethical dimensions of disease classification, and so naturalistic accounts are clearly non-starters from the standpoint of a conceptual role analysis that takes a broader view of the relevant institutions. Naturalism, especially if it is indexed to evolutionary history, may be capable of providing a unified definition of disease—but any such an account would be biological, not medical (more on this below). A defender of disease naturalism might argue that the ethical-institutional dimensions of the disease concept are actually best served by attempting so far as possible to keep moral values out of disease classification, given the historical abuses that have flowed from value infiltration. But if so, then the specification of ‘disease’ (if not individual diseases) would be justified by reference to explicitly moral values—yet naturalists not defend their specification of ‘disease’ in this way, and it is unclear whether such an approach could properly be called naturalistic at all. Moreover, a key motivation for moving away from naturalism, as we have seen, is that value-free disease classification has never been achieved in practice,⁵⁰ nor would such an achievement be desirable given the population ethical desiderata discussed above.

Purely normativist approaches to disease fare no better, even if they were amended to be thickly normative. On the one hand, to dispense with a biological component *entirely* would be to classify all properly disvalued traits as diseases—even ‘purely social’ ones—and this would gut the epistemic value of the disease concept. On the other hand, to dispense with the biological *dysfunction* component would mean classifying as a disease any biomedical state of affairs that is disvalued (or would be properly disvalued). The problem with that, as we have seen, is that it does not allow the disease concept to prioritise biomedical conditions that tend to interfere most with well-being, flourishing and opportunity. There will be exceptions, of course, and difficult coverage decisions will have to be made. For instance, to borrow an example from Buchanan *et al.*,⁴⁷ if Johnny is expected to be shorter-than-average due to a hormone deficiency, whereas Jimmy is expected to be shorter-than-average because he has short parents, do we have sound ethical reasons for prioritising Jimmy’s ‘treatment’ over Johnny’s ‘enhancement’, even though the effects on well-being and opportunity are the same? A purely normativist account of disease has no basis by which to distinguish these cases. The point is not that Johnny-like cases should be prioritised over Jimmy-like cases; the point is that a thickly normative hybrid view explains why Jimmy’s case is a hard case, providing us with the starting point for identifying when, where and for what reasons exceptions to the treatment/enhancement distinction might be warranted.

The weakly normative hybrid account performs better in many regards than naturalist and normativist theories. It includes a biofunction component that restricts its range of application and thus plays a useful role in healthcare prioritisation; and it affords a central place for moral evaluation in disease classification. However, the weakly normative hybrid view ultimately fails to account for misapplications caused by wrongful evaluations, and it falls back into the spatiotemporal contingency problems that beleaguer weakly normative normativist accounts.

5. ELIMINATING ELIMINATIVISM

The foregoing considerations counsel against jettisoning the disease concept in medical contexts, as some theorists have recommended we do.¹ Disease eliminativism argues that given the problems with ‘disease’ both in concept and in application, we ought to adjust our healthcare environment to function in a ‘disease’-neutral setting—speaking only of ‘biomedical states that are disvalued’. The welfarist approach to disease (§3) arrives at essentially the same conclusion, although for different reasons (namely, the purported moral arbitrariness of the treatment-enhancement distinction). The problem with eliminativism, whether of the standard or welfarist variety, is that it is unclear how avoiding talk of disease resolves the problems discussed above.

If eliminating the disease concept involved no more than expunging the word from our healthcare vernacular, at best this would be an empty terminological victory with the replacement term, possessing the same social-informational significance, emerging to take its place. If, on the other hand, eliminating the disease concept involved something more substantive than this, then it is not clear that such an elimination would be desirable. For as we have seen, the disease concept is interwoven into the institutional fabric of developed societies and there are population-ethical justifications for its embeddedness that are not satisfied by this replacement term. Because standard eliminativism lacks a proper evaluation component, it fails to protect against the risks associated with misapplications and abuses of the concept. And because both standard and welfarist forms of eliminativism lack a biofunctional component, they do not serve the ethical goals of healthcare allocation and prioritisation.

It is true that, by distinguishing biomedical state descriptions from value judgements, the standard eliminativist approach guards against the unwitting infiltration of values into classification schemes that are pedalled as naturalistic descriptions. The scientific labelling of certain conditions as biological dysfunctions can no doubt have troubling social consequences, particularly given the (unfortunate) value-laden connotations of the natural/unnatural distinction (§4). But what drives pernicious misapplications of ‘disease’ is not so much the smuggling of value into supposedly descriptive classifications—rather, it is the *improper evaluation of biomedical states*. As such, even if eliminating the disease concept were possible, doing so would fail to address the ethical problems that motivate eliminativism in the first place.

Finally, eliminating the disease concept also fails to capture some of the important sociocultural sequelae that flow from disease diagnosis. For instance, in 2008 the Canadian Medical Association⁵¹ found that nearly half of all Canadians believed that some things described as mental illness were no more than ‘an excuse (for a person’s) poor behaviour and personal failings’ (p. 26). Glannon⁵² rightly points out that biological psychology has done much in helping to dispel pernicious views that mental illnesses are ‘all in the mind’ or somehow less ‘real’ than physical conditions (p. 1). Coming to regard mental illnesses (such as PTSD—see box 1) as medically treatable pathologies, rather than as weaknesses of character or fortitude, has greatly reduced their social stigma and the moral marginalisation of those who suffer from them.

Everything we have said is compatible with the eliminativist claim that we have good reason to distinguish biomedical states from our evaluations of those states. But, as we have seen, this is already achieved by the hybrid account, and it is at least implicit in the naturalistic account. Moreover, both clear-cut

cases and controversial borderline cases are better managed by the proposed view than they are by existing theories of disease. Take, for instance, an uncontroversial case like contraception: on naturalistic accounts (whether grounded in Boorsean normal function or selected effects function), birth control pills induce a disease in patients, namely premature (if reversible) infertility. In contrast, the proposed hybrid view does not consider medically induced reproductive dysfunction to be a disease, since the state is properly valued (ie, it contributes to well-being, flourishing, opportunity and autonomy without doing harm to others). By the same token, if the same reproduction-disabling hormonal state occurs naturally in a patient, then that state is properly considered a disease on the present view since unchosen infertility is rightfully devalued at the population level. Whereas naturalistic accounts run into difficulties with properly valued (or valuable) dysfunctions, our proposed account has little trouble excluding such cases from the category of disease.

What about more controversial conditions? On our proposed account, controversial cases will continue to be controversial whenever they do not clearly implicate a bio-dysfunction or when it is unclear whether the underlying dysfunction ought to be disvalued: two problems that are especially pronounced in the domain of mental illness (box 1). In such cases, the disease label ought to be used cautiously and final judgements suspended; but as with all categories, we need not conclude from the existence of hard cases that the disease concept ought to be abandoned.

CONCLUSION

Let us now take stock. The analysis of concepts, whether moral, scientific or (as in the case of 'disease') a combination of the two, cannot take place in an institutional vacuum. Concepts are specified in relation to institutions and are shaped by particular pragmatic, epistemic or ethical goals. We have argued that naturalistic theories of disease do not succeed, not because they fail on their own terms, but because they do not suit the role the concept currently plays in our healthcare institutions. Normativist and thinly normative hybrid theories fail to perform much better in this regard. Although existing accounts of disease suffer from various flaws, eliminativism likewise fails to appreciate the important role the concept currently plays in human healthcare. Thus, we have defended an amended hybrid theory that is objective through and through. In particular, it takes on two burdens of objectivity: biological objectivity, which requires that we reliably identify dysfunction, and moral objectivity, which requires that our reasons for disvaluing a given dysfunction are subject to rational justification. Both of these tasks pose formidable conceptual and methodological challenges that warrant further discussion.

In closing, we stress that we do not intend that this argument extend to other sciences that make their own fruitful use of the disease concept. Theoretical unification is a worthwhile scientific pursuit, and since human medicine may reasonably be viewed as a subset of biological science, one might argue that we should aim for concepts that unify the medical and biological domains. Yet a concept of disease that is useful in, say, veterinary medicine or forestry science may be poorly suited to the thickly normative aims of human medicine. Furthermore, the moral institutional dimensions of the disease concept are not limited to matters of healthcare allocation. Classifying certain conditions as diseases could have implications in other institutional contexts, such as in criminal and tort law where at issue is the apportionment of responsibility for purposes of assigning punishment, rather than the allocation of medical resources. It is not clear that merely

classifying some mental condition as a disease in the sense we have proposed should have any general implications for the moral evaluation of mental states or the voluntariness of action. In any case, our proposal is restricted to the healthcare context and we make no pretensions about how the disease concept ought to be used in other institutional settings.

We do not advocate a pluralistic approach to disease within the broad domain of human healthcare, as Boorse seems to do in his attempt to distinguish theoretical from clinical notions of disease. This is because the institutional aspects recounted above speak strongly in favour of conceptual unity in this domain. However, the foregoing discussion also shows that a truly procrustean approach to 'disease', such as one based on biological dysfunction alone, is undesirable in the context of human healthcare. A similar argument could be extended to debates over medical definitions of 'death',^{53–55} where the epistemic benefits of a universalisable definition must be weighed against the ethical costs that flow from undercutting the institutional role the concept of death plays in human healthcare, most notably in relation to organ procurement and the lost opportunity costs that come from keeping a dead person's body alive in perpetuity.

We do not expect this to be the last word on 'disease', and we anticipate that some will find our proposal wanting or incomplete. However, we hope to have made a convincing case for at least two basic claims. First, that the frequent diametric opposition of naturalist/objectivist and value-based/subjectivist accounts is in fact a false dichotomy; and second, that the disease concept, however it is specified, ought to be tailored to the role that the concept plays in the institutional settings in which it is deployed. If these two conclusions are accepted, then the 'disease' debate takes on clearer rules of adjudication, the space of plausible theoretical options narrows considerably and significant progress can be made towards a resolution. Incorporating proper evaluation into disease classification is to acknowledge that medicine is a robustly and irreducibly normative discipline. Some authors and practitioners may find this disquieting, but we suspect many will welcome moral values into the practice of medicine, and take comfort in knowing that values have always been at the very heart of the discipline.

Acknowledgements The authors would like to thank Marc Ereshefsky, Allen Buchanan, Şerife Tekin, Sune Holm, Steve Clarke, two anonymous referees, as well as audiences at the PSA 2018, University of Miami, University of Oxford, University at Albany, Rochester Institute of Technology, and Boston University, for comments on an earlier draft of this manuscript.

Contributors The authors contributed equally to this paper.

Funding Scarffe is grateful to the Social Sciences and Humanities Research Council of Canada for support of this research.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

- 1 Ereshefsky M. Defining 'health' and 'disease'. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 2009;40:221–7.
- 2 Walker MJ, Rogers WA. Introduction: the boundaries of disease. *J Med Philos* 2017;42:343–9.
- 3 Brigandt I. The dynamics of scientific concepts: the relevance of epistemic aims and values. In: Feest U, Gruyter D, Steinle F, *Scientific concepts and investigative practice*, 2012:75–103.
- 4 Bayer R. *Homosexuality and American Psychiatry: the politics of diagnosis*. Basic Books, 1981.
- 5 Spitzer RL. Harmful dysfunction and the DSM definition of mental disorder. *J Abnorm Psychol* 1999;108:430–2.

- 6 Tekin S. Psychiatric taxonomy: at the crossroads of science and ethics. *J Med Ethics* 2014;40:513–4.
- 7 Bingham R, Banner N. The definition of mental disorder: evolving but dysfunctional? *J Med Ethics* 2014;40:537–42.
- 8 Engelhardt HT. The disease of masturbation: values and the concept of disease. *Bull Hist Med* 1974;48:234–348.
- 9 Boorse C. Health as a theoretical concept. *Philos Sci* 1977;44:542–73.
- 10 Boorse C. In: Humber M, Almeder RF, "A rebuttal on health," in *What is disease?* Humana Press, 1997:3–143.
- 11 Hull DL. On Human Nature. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* 1986;1986:3–13.
- 12 Amundson R. Against normal function. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 2000;31:33–53.
- 13 Lewens T. Human nature: the very idea. *Philos Technol* 2012;25:459–74.
- 14 Sisti D, Young M, Caplan A. Defining mental illnesses: can values and objectivity get along? *BMC Psychiatry* 2013;13:346.
- 15 Powell R, Scarffe E. Rehabilitating 'disease': function, value, and objectivity in medicine. *Philosophy of Science* 2019;86.
- 16 Millikan RG. *Language, thought, and other biological categories*: MIT Press, 1984.
- 17 Neander K. Functions as selected effects: the conceptual analyst's defense. *Philos Sci* 1991;58:168–84.
- 18 Engelhardt HT. Ideology and Etiology. *Journal of Medicine and Philosophy* 1976;1:256–68.
- 19 Margolis J. The concept of disease. *Journal of Medicine and Philosophy* 1976;1:238–55.
- 20 Goossens WK. Values, health, and medicine. *Philosophy of Science* 1980;47:100–15.
- 21 Sedgwick P. *Psycho Politics*: Harper & Row, 1980.
- 22 Fulford KW, Smirnov AY, Snow E. Concepts of disease and the abuse of psychiatry in the USSR. *Br J Psychiatry* 1993;162:801–10.
- 23 van Voren R. Political abuse of psychiatry-an historical overview. *Schizophr Bull* 2010;36:33–5.
- 24 Szasz TS. The myth of mental illness. *Am Psychol* 1960;15:113–8.
- 25 Wakefield JC. Disorder as harmful dysfunction: a conceptual critique of DSM-III-R's definition of mental disorder. *Psychol Rev* 1992;99:232–47.
- 26 Basl J. *The death of the ethic of life*. Oxford: Oxford University Press, 2018.
- 27 Cooper R. Disease. *Studies in history and philosophy of biological and biomedical sciences*. 2002;33:263–82.
- 28 Wakefield JC. The concept of mental disorder: diagnostic implications of the harmful dysfunction analysis. *World Psychiatry* 2007;6:149–56.
- 29 Kingma E. Naturalism about health and disease: adding nuance for progress. *J Med Philos* 2014;39:590–608.
- 30 Crasnow S. Feminist philosophy of science: 'standpoint' and knowledge. *Sci Educ* 2008;17:1089–110.
- 31 Blackburn S. In: Timmons M, Sinnott-Armstrong W, "Securing the notes: moral epistemology for the quasi-realist," in *Moral knowledge?: new readings in moral epistemology*. Oxford : Oxford University Press, 1996.
- 32 Street S. A darwinian dilemma for realist theories of value. *Philos Stud* 2006;127:109–66.
- 33 Railton P. Moral realism. *The Philosophical Review* 1986;95:163–207.
- 34 Dearborn LW. The problem of masturbation. *Marriage Fam Living* 1952;14:46–55.
- 35 Pinker S. The stupidity of dignity. *The New Republic* 2008.
- 36 Macklin R. Dignity is a useless concept. *BMJ* 2003;327:1419–20.
- 37 Powell R. The law and philosophy of preventive war: an institution-based approach to collective self-defense. *Australian Journal of Legal Philosophy* 2007;32:67–89.
- 38 Buchanan A. Institutionalizing the Just War. *Ethics and International Affairs* 2005;34:2–38.
- 39 Brock DW. Separate spheres and indirect benefits. *Cost Eff Resour Alloc* 2003;1:4.
- 40 Roache R, Savulescu J. Psychological disadvantage and a welfarist approach to psychiatry: an alternative to the dsm paradigm. *Philosophy, Psychiatry and Psychology* 2018;25:245–59.
- 41 Fulford KW. Values-based practice: Fulford's dangerous idea. *J Eval Clin Pract* 2013;19:537–46.
- 42 Tekin S. "Against hyponarrating grief: Incompatible research and treatment interests in the DSM-5," in *The DSM-5 in Perspective*: Springer, 2015:179–97.
- 43 Tekin S, Mosko M. Hyponarrativity and context-specific limitations of the DSM-5. *Public Affairs Quarterly*. forthcoming.
- 44 Stramondo JA. The distinction between curative and assistive technology. *Sci Eng Ethics* 2018.
- 45 McMahan J. The morality of screening for disability. *Reprod Biomed Online* 2005;10:129–32.
- 46 Wakefield JC. The biostatistical theory versus the harmful dysfunction analysis, part 1: is part-dysfunction a sufficient condition for medical disorder? *J Med Philos* 2014;39:648–82.
- 47 Buchanan A, et al. *From chance to choice: genetics and justice*: Cambridge University Press, 2001.
- 48 Harris J, Chan S. Enhancement is good for you!: understanding the ethics of genetic enhancement. *Gene Ther* 2008;15:338–9.
- 49 Williams GC, Nesse RM. The dawn of darwinian medicine. *Q Rev Biol* 1991;66:1–22.
- 50 Plutynski A. Ethical issues in cancer screening and prevention. *J Med Philos* 2012;37:310–23.
- 51 Canadian Medical Association. 8th annual national report card on health care. 2008.
- 52 Glannon W. The blessing and burden of biological psychiatry. *Journal of Ethics in Mental Health* 2008;3:1–4.
- 53 Bernat JL. A defense of the whole-brain concept of death. *Hastings Cent Rep* 1998;28:14–23.
- 54 Shewmon AD. The brain and somatic integration: insights into the standard biological rationale for equating "brain death" with death. *J Med Philos* 2001;26:457–78.
- 55 Veatch RM. The impending collapse of the whole-brain definition of death. *Hastings Cent Rep* 1993;23:18–24.