Journal of Medicine and Philosophy, 45: 145–158, 2020
doi:10.1093/jmp/jhz042
Advance Access publicaton on January 10, 2020

Health and Reference Classes

SANDER WERKHOVEN*
Universiteit Utrecht, Utrecht, The Netherlands

*Address correspondence to: Sander Werkhoven, PhD, Universiteit Utrecht, Janskerkhof 13, Room 1.06, 3512 BL Utrecht, The Netherlands. E-mail: s.werkhoven@uu.nl

In this article, I address two objections developed by Kingma against Boorse's (1977) bio-statistical theory of health, the objections that choice of reference classes renders the theory both circular and problematically value-laden. These objections not only apply to the bio-statistical theory of health but also to other naturalistic theories, like the dispositional theory of health. I present three rejoinders. First, I argue that the circularity objection arises from excessive methodological demands. Second, I argue that naturalists can resist the normativist claim that health and pathology are differentiated on the basis of personal or cultural values. Finally, I show that it is possible to justify choices between rival theories of health without the interference of evaluative commitments. With these rejoinders, I conclude that the bio-statistical theory, as well as other naturalistic theories of health utilizing reference classes, is not undermined by Kingma's arguments.

Keywords: *health, naturalism, normativism, reference class*

I. INTRODUCTION

In this article, I address an argument put forward by Kingma (2007) against the bio-statistical theory of health (BST) defended by Boorse (1975, 1976, 1977, 1997, 2014). According to Boorse, health is a state of normal functional ability; that is, an organism is healthy if all bodily parts are able to make a statistically typical contribution to the organism's overall goals of survival and reproduction. What counts as "normal" or "statistically typical," on this account, depends on the reference class in which normality is determined: a relatively uniform class of organisms specified on the basis of species, sex, and age. Hence, on the BST an organism is healthy if all its functional parts are able to function in a statistically normal way, given its species, sex, and age.

The main appeal of the BST is that it purports to be a naturalistic theory of health and pathology, free from people's subjective values and culturally accepted norms about how a person should be. Kingma's (2007) objections attack the BST's proclaimed value-neutrality. She argues that Boorse's use of reference classes renders the view dependent on subjective values in a way the BST aimed to eschew, while simultaneously rendering the account problematically circular. If the BST's proclaimed naturalism proves untenable in light of Kingma's arguments, this would be a significant blow to the naturalist camp in the philosophy of medicine. There are other theories of health and pathology that utilize the idea of reference classes, which equally claim to be naturalistic.[1] If Kingma's objections hold, all such theories would be undermined—at least in their ambition to offer a value-free, naturalistic account of health and pathology.

Kingma's critique has not yet been met with a satisfactory response. Boorse's (2014, 964) brief treatment of the objection falls short of being a convincing reply. My aim in this article is to refute Kingma's objections and to defend the deployment of reference classes in naturalist theories of health. To that end, in Section II, I bring Kingma's argument into focus by presenting it in the context of its original target. In Section III, I briefly describe Boorse's response and highlight its shortcomings. In Sections IV–VI, I take on Kingma's objections, first tackling the circularity claim and subsequently the value-ladenness concern.

II. KINGMA'S OBJECTIONS TO THE BST

In a series of articles from the 1970s, Boorse developed a theory of health and pathology with the aim of capturing how these concepts are used in medical theory.[2] In theoretical medical discourse, Boorse claims, pathology is standardly equated with abnormal functional abilities of functional parts and good health with normal functional ability of functional parts (1977, 546). Boorse endorsed and systematized this use into a theory of health and pathology: when functional parts like a heart, kidney, respiratory system, and eye are able to function within margins of normal variation, one is healthy; when they function below these margins, one is in a pathological state. The function of a part, on Boorse's view, depends on the contribution that part makes to the organism's survival and reproductive abilities—the overall goals of the organism (1977, 556). What counts as "normal" functioning is to be understood in strictly statistical terms according to Boorse, although he never specifies how far efficiency outcomes have to drop below average efficiency levels for them to qualify as dysfunctional forms of functioning.

Boorse explains at length that statistically normal functioning is to be determined in a group sharing a "coherent functional design" (1977, 558, 1997,

37–41, 2014, 694–91). To know whether a heart functions normally, its efficiency has to be compared with the average efficiency of hearts of organisms belonging to a "natural class of organisms of uniform functional design" (Boorse, 1977, 555). This uniform group of organisms, in which normal functional ability is determined, is called a "reference class" and is constituted by organisms of similar species, sex, and age, according to Boorse. The BST needs these reference classes, as Kingma rightly emphasizes, since "the human species shows a wide variety of functioning; what is normal in one group can be abnormal in another" (2007, 128). An adult human being with the communicative abilities of a 2-year-old toddler, for instance, will generally be considered pathological. The BST accounts for this: in the reference class specified by species, sex, and age, this adult indeed displays abnormal functional abilities. Hence, whether you are healthy or pathological on the BST depends on whether your functional parts can function in ways that are statistically normal for the reference class of (fe)male human beings of roughly your age.

What Boorse offers is a reductive analysis of the concepts health and pathology: the nature of health and pathology is demystified and brought back to the base level of part-functioning ability and statistically normal efficiency outcomes within a given reference class. In doing so, Boorse believes to have put forward an accurate, naturalistic, value-free conception of health: whether one is healthy or pathological does not depend on personal values, on broader cultural norms of how one should be, on what a person wants to do or is accustomed to doing, or on any other type of evaluative attitude toward one's condition.[3] Boorse (1997, 2014) has remained adamant that the BST relies on empirical facts only, not evaluative judgment.

Kingma's (2007) critique of the BST centers on Boorse's use of reference classes. If reference classes were chosen differently from those stipulated by Boorse, the dividing line between health and pathology would also fall differently. If myopia would constitute a further division of reference classes, for instance, being short-sighted would be statistically normal within the reference class of short-sighted people and therefore not qualify as a pathology. Likewise, Kingma argues, if Down's syndrome were used as a reference class, people with Down's syndrome would display normal functional ability in this reference class and therefore count as normal, that is, healthy (2013, 40).

The fixing of reference classes is therefore of critical importance to the BST. Because of its importance, Kingma argues that Boorse "needs to give an account of the distinction between reference classes that are allowed and reference classes that are not allowed" (2007, 129). In a later summary of the argument, she formulates the demand slightly differently: she says Boorse has to "provide a value-free justification" for what may pass as an appropriate reference class and what may not in judgments about health (Kingma, 2013, 40). The natural response to this challenge would be to insist

that conditions like myopia and Down's syndrome ought not to constitute a reference class because these conditions *just are* pathological conditions. However, this response is unavailable to Boorse, Kingma is quick to point out, for if decisions about the appropriateness of reference classes can only be justified on the basis of prior knowledge about the distinction between health and pathology, "then his account is circular" (2007, 129). The first objection, then, is that choice and justification of reference classes on the basis of prior knowledge about who counts as healthy and pathological renders the BST circular.

Kingma further points out that the notions Boorse invokes to justify his choice of reference classes—naturalness, uniformity, and design—do not help separate appropriate from inappropriate reference classes (2007, 129–31). Down's syndrome is as natural, uniform, and genetically "designed" as being male. From this Kingma draws the conclusion that Boorse's choice of reference classes is not rooted in empirical facts, but, instead, derives from "an evaluative choice which may reflect some deep underlying normative commitments" (2007, 132).

To add further support to this conclusion, she imagines two theories of health, Boorse's BST and XST, whereby XST is identical in all respects to the BST except for having one further specification of reference class: sexual orientation (Kingma, 2007, 132). On the BST, homosexuality is abnormal with respect to reproductive functioning and thus a pathology; on the XST, homosexuality is statistically normal in the reference class of homosexual people, hence not a pathology. If we now ask whether homosexuality is a disease or not, we are in effect asking whether BST or XST is the correct theory of health. Since at this point there is no relevant empirical data that directs us one way or the other, Kingma concludes only our prior evaluative commitments can inform the decision. This, then, leads to the second objection: fixing reference classes is based on prior evaluative judgments about who is to count as healthy and as pathological. Hence, the BST is not value-free, it only formalizes antecedently held evaluative judgments about certain conditions. As independence from values is the main attraction of the BST, Kingma believes this objection leaves the BST "thoroughly, and fatally, undermined" (2013, 40).

III. BOORSE'S RESPONSE

Boorse (2014, 694) has responded to Kingma's objection with an argument by analogy—the analogy being the definition of the concept "mare." On one theory "mare" is defined as "adult female horse" (HT), on another "mare" is defined as "adult female mammal" (MT). Boorse thinks that the question whether HT or MT is correct is analogous to the question whether BST or XST is correct. HT and MT, Boorse notes, "differ only in their basic

reference class: horses vs. mammals" (2014, 694). Take an adult female pig, ask whether it is a mare, and we will have to choose whether HT or MT is the correct account of marehood, just like asking whether homosexuality is a disease forces us to choose between BST or XST. In the pig case, the answer is of course that HT is correct and that the pig is not a mare. The crucial point this argument is supposed to demonstrate, however, is that there are no immediately evident *values* involved in settling this question. Likewise, when asking whether homosexuality is a disease and so whether BST or XST is correct, Boorse thinks we can turn to professional medical usage of the terms and conclude that BST is correct—and crucially, this does not reveal, or depend on, prior subjective or cultural valuing or disvaluing certain conditions.

This argument suffers from two weaknesses. First, the analogy does not quite work: the HT/MT example does not involve reference classes in the same way that BST and XST do. On the BST, one is healthy if and only if all part-functional abilities are not statistically subnormal. Statistical normality, as we have seen, is to be determined in the appropriate reference class. Reference classes in the BST, therefore, determine the very meaning of "statistical normality": what "statistical normality" *is* depends on the reference class in which it is determined. On the HT, as Boorse sees it, an animal is a mare if it is an adult female in the reference class "horse." The reference class "horse," however, says nothing about the meaning of "female": what "being female" *is* does *not* depend on the reference class (horse) in which it is determined. Hence, reference classes in the BST/XST case play a different role compared with Boorse's HT/MT analogy. It is not immediately clear whether this disanalogy really undermines the success of the argument, but it must be noted that in MT and HT the idea of a reference class is used quite differently from the way it is used in BST and XST.

Second, the force of Boorse's argument can be attributed to the analyticity of the claim that a mare is an adult female horse. An audience with a sufficient command of English learns nothing about mares when told that mares are adult female horses. Reductive accounts, by contrast, typically tell us something we previously did not know: if the BST is correct, we would learn that the difference between health and pathology turns out to be a difference between statistically normal and abnormal functioning of organisms' parts. The analyticity of HT's truth in the case of the mare analogy is important because it sidesteps the test that Kingma pressed. When asking whether BST or XST is correct, we cannot rely on analytic truths: it is not true by definition that homosexuality or myopia is healthy or pathological—such knowledge is not possessed merely through understanding the relevant terms.

This last point may be resisted. One could argue that the argument does not rest on an appeal to analytic truths; instead, knowledge of whether HT or MT is correct depends on how biologists use the concept "mare," and this is the only input required to decide that HT corresponds best to this

use—hence, no analyticity at work here. Likewise, when we want to know whether BST or any XST-variant is correct, we should consult medical professionals and see whom they count as healthy and unhealthy. With this type of input, we may conclude whether BST fits the expert belief about instances of health and pathology better than any XST-variant.

Will this convince Kingma and those persuaded by her arguments that the BST does not rely on values? Probably not. Critics can easily maintain that if judgments by experts are indeed best captured by BST instead of some XST-variant, this only goes to show that value judgments of a certain professional body at a certain time are best captured by the BST. Had the evaluative commitments of medical professionals been different, some XST-variant might have won the contest for best fit with professional opinion. Kingma's conviction that choice between competing theories of health turns on values would remain unscathed.

Boorse's argument nevertheless helps to gain some ground on critics like Kingma. The argument shows that choices between two different theories (including theories only differing in reference classes) of some fundamental concept *need not* depend on evaluative judgment in the way that Kingma immediately assumes. What the argument does not achieve, however, is convincingly show that the choice of theory, and thus of reference classes, *does not* depend on evaluative commitments shared by medical professionals or some other group. To argue this stronger point, and also to dismiss the circularity concern, is what I shall turn to now.

IV. THE CIRCULARITY CONCERN

Kingma's first objection is that theories of health are circular if they employ reference classes that can only be justified on the basis of prior knowledge about the health/pathology distinction. This rests on a mistake. The mistake is to confuse the *justification* of a given definition with the *meaning* of the terms used in that definition. For an audience to understand what is meant by "statistically normal functional ability of parts within a reference class," it is not required they already know the meaning of the concepts health and pathology; if one inquires further into what is meant by "statistical normality," "functioning of parts," or "a reference class," perfectly intelligible answers can be given that do not presuppose knowledge of the health/pathology distinction.

What Kingma appears to demand, if circularity is to be avoided, is that the *justification* for adopting a theory of health does not rely on prior ideas about who is to count as healthy and who as pathological. As a methodological demand, this is too stringent. To justify the choice of a philosophical theory, one may refer to all sorts of antecedently held beliefs about the phenomenon in question. A justification for a certain theory might be that it fits

best with common sense intuitions or that it succeeds in accommodating and explaining certain paradigm cases of healthiness or pathology. Naturally, intuitions or paradigm cases involve prior understanding of and beliefs about the phenomenon, but in a *justificatory* role utilizing these prior beliefs is not problematic—this is how philosophical analyses generally proceed. Take a metaphysician working out a theory of causality: of course, she will rely on cases of certain *prima facie* cause-effect relations that must be explained by her theory. If the theory captures them, it gives credence to the theory and justifies favoring this theory over rival ones. Similarly, if a moral philosopher claims her theory has a better fit with intuitive plausible cases of right action than rival theories, this will not render her theorizing circular (because it "presupposes" knowledge about right action rather than "yielding" it). In Boorse's case, it is not everyday intuitions that guide the analysis, of course, but views shared by medical professionals. The recognition that fixing appropriate reference classes requires prior knowledge about cases of healthy and pathological neither stems from methodological licentiousness nor results in definitional circularity.

V. JUSTIFICATION OF REFERENCE CLASSES

This point brings us to the question whether the justification of reference classes depends on subjective values and cultural norms. Kingma's value-ladenness objection states that choosing reference classes ultimately relies on an evaluative choice, because a strictly empirical justification of choosing reference classes is unavailable. To see what goes wrong in this argument requires some unpacking. This is what I take Kingma's argument to be[4]:

(i) pathological conditions and healthy conditions are differentiated and identified on the basis of conditions that we (or some group) subjectively value and disvalue, that is, on the basis of cultural norms about how people ought to be[5];
(ii) after this initial sorting, we may try to give a naturalistic description of these two groups in theories like Boorse's BST[6];
(iii) but crucially, we choose and adjust reference classes in theories like the BST so that our evaluative judgments made in (i) remain intact and decisive;
(iv) only the initial evaluative choices made in (i) can justify the choice of reference class and the specific version of the BST that its author endorses;
(v) therefore, both the BST (and all other naturalistic theories employing reference classes) are essentially value-laden.

It should first be noted that it is possible to accept (i)–(iv) and still reject conclusion (v). Naturalists may concede that the initial sorting of pathological

versus healthy conditions occurs on the basis of values held by particular individuals or communities. Once this sorting has been accomplished, we then investigate whether the conditions that have been dubbed pathological share features other than "being disvalued"—features that are simultaneously specific to this group. If there are not any, then the concepts health and pathology prove to be irreducibly value-laden—they *only* reflect our evaluative appraisals of certain physical and mental states. If there are such shared features, one may attempt a reduction of the original value-laden concepts to the level of the shared base-level descriptive features and assess the accuracy and specificity of the reductive view. Contra (v), however, the resulting reductive theories are not necessarily value-laden; if the reduction succeeds, the evaluative concepts will have been reduced to descriptive concepts. So the simple response to the argument stated in (i)–(v) is that the initial sorting and ultimate justifications for categorizing conditions as healthy or pathological may indeed be based on subjective or cultural values, but the BST and other naturalistic theories of health aim to offer a reduction of the concepts at the level of value-free physical description and may very well succeed in doing so.[7] I take it that Kingma would concede this response, but that she would insist that values will have played a decisive role and therefore deny that the naturalist has truly offered a value-free account of health and pathology. After all, values intruded in what she calls "the operationalization of function and dysfunction" (Kingma, 2014, 594). What she means by this is that in the operation of defining health and pathology, evaluative judgments have played a decisive role, even though the result may be a theory formulated entirely in naturalistic terms.

Now a stronger line of defence is also available to the naturalist, one that shows more robustly that Kingma's value-ladenness concern is unconvincing. Naturalists may also deny (i) and, on that basis, reject the rest of the picture, including the crucially important claim made in (iv). The claim made in (i) implies that all pathological conditions have at least one feature in common: being negatively valued by us (or at least some of us)—and correspondingly, that all healthy conditions share the feature of not being negatively valued by us (or at least some of us). Naturalists like Boorse, however, have good reasons to reject (i) from the outset. First, it is perfectly possible for individuals or cultures to value positively clear cases of pathologies and to value negatively clear cases of health, without this altering the health-status of these conditions.[8] Boorse mentions the example of cowpox, which could be hugely valuable during an epidemic of smallpox, and sterility, which in a time without contraception may make for a true blessing if one already has a large family (Boorse, 1977, 545). The positive value of cowpox and sterility certainly would not stop them from being pathologies. The dependence on individual or cultural values implied by (i) does not mesh well with the apparent objectivity of a large class of pathologies.[9] Second, there are many physical and mental conditions we tend to disvalue,

both personally and culturally, that are neither diseases nor health impairments—ugliness, shortness, and stupidity being cases in point.[10] If (i) were true and the initial sorting of health and pathological conditions occurs on the basis of values, it is likely these conditions would come out equally as pathologies. Third, the conditions picked out by the concepts health and pathology share nonevaluative features that explain their commonality much better and more reliably. When analyzing the conditions that we generally conceive of as healthy and pathological, we discover that what underlies the phenomena is not our evaluative attitudes toward this or that physical and mental condition, but the objective features pointed to by the BST (or a better naturalistic theory). Evaluative attitudes may most of the time correspond to this reality, but these three reasons indicate they are not constitutive of health and pathology in the way that Kingma claims.

To insist, then, as Kingma does with (iv), that a choice of reference class can only be justified on the basis of prior evaluative commitments means obscuring the fact that features other than our values underpin the distinction between health and pathological conditions. Health and pathology, the naturalist can maintain, are properties of living organisms that exist independently from our evaluative judgments. Although we may *prima facie* get to know the phenomena through our evaluative attitudes toward (un) healthy conditions, when we analyze them, we discover that what underlies the distinction between health and pathology is just what the BST claims it is (or what a better rival naturalistic theory claims it is). This is the reality to which our evaluations turn out roughly to *correspond*, rather than *give rise*. When Kingma asks for a value-free justification for choosing certain specific reference classes in defining health and pathology, the naturalist can point directly to the objectivity of health and pathology. Because health and pathology are not initially differentiated on the basis of evaluations, it is also not evaluative judgments that justify the choice of reference classes; what justifies the choice of species, sex, and age as the relevant reference classes is that health and pathology—*qua* objective phenomena—can only be adequately captured and defined with the help of these reference classes, and not with any other.

VI. JUSTIFICATION FOR CHOICE OF THEORY

Besides rejecting the crucial premise on which Kingma's argument hinges, (i) above, there is another strategy to undermine her objection. Recall that she pressed her point by imagining two theories, BST and XST, where XST is similar in all respects to BST except for using medical condition *X* as a further specification of reference classes. Asking whether *X* is a disease means choosing whether BST or XST is the correct theory of health, which, according to Kingma, can only be settled on the basis of our evaluative attitudes toward

X.[11] In the face of two competing theories of health with different reference classes—BST versus some XST-variant—how can we choose one theory over another without our evaluative commitments guiding the decision-making process? Pointing out, as I have done in the previous section, that we have good reasons to think that health and pathology are properties of living organisms that exist independently from our evaluative judgments—so that it is not our values that determine whether BST or any XST-variant is correct, but the objective reality of health and pathology—does not really help when we are actually confronted with a BST/XST choice.[12]

We should still question, though, why it would be specifically our *values* that would make us choose one of any two competing theories of health. In all domains of philosophy, a successful theory has to meet various requirements and is not accepted just because it sits well with our prior evaluative commitments about individual cases. Criteria like fit-with-paradigm cases, theoretical specificity, congruence with other accepted (or plausible) philosophical and scientific theories, ontological parsimony, epistemic access to the phenomenon in question, and so on, all constitute good reasons to favor one theory over another. Why then, in the case of health, would our personal or cultural values about who is to count as healthy be the deciding factor in choosing between rival theories? It is clear that the BST is simpler than all XST-variations (at least one less factor determining the reference class); more uniform (it does not deploy *ad hoc* conditions like homosexuality or myopia as specification of reference classes); and it has a better fit with paradigm cases (it will label myopia, Down's syndrome, and all other medical conditions introduced as reference class in XST-variants as a pathologies). Whether we value or disvalue the condition used to formulate an XST-variant therefore need not play any role in deciding that the XST-variant is a worse candidate for a theory of health.

But still, these criteria may not always make it clear which theory is best, especially when it concerns a choice between two theories that rule differently over just one contentious case (like in Kingma's example). At this point we cannot rely on fit with paradigm cases—the only differing outcome is, after all, a contentious condition—and relying solely on one theory being slightly simpler than another also is not epistemically conclusive. So, how can this challenge be met without falling prey to value judgments? One way to approach this issue is by reflecting on what it means for a condition to be "contentious" with respect to health. If it is not clear whether the condition is to count as a pathology, this means people with the condition have to be compared to people without it. If we want to know whether myopia is a pathology, we have to compare myopic people with non-myopic people and assess (if the BST is correct) whether the various optical organs are functioning statistically normally or abnormally. We can carry out such an assessment only if we contrast myopic people to clear-sighted people. This *already* implies that myopia cannot serve as a specification of reference

classes in definitions of health. For if myopia were to serve as specification of a reference class, we would end up comparing myopic people to other myopic people and find out nothing about their health status. What I am suggesting, then, is that we have an intuitive grip on the sort of conditions that *could* be healthy or pathological, and these conditions cannot constitute a reference class. If it is germane to ask whether some condition *C* is a pathology, this means it must be possible to determine whether *C* is a pathology by comparing organisms with *C* to organisms without *C*, which means that the condition *C* cannot function as a reference class. So in Kingma's example: if it is germane to ask whether homosexuality is healthy or pathological, this means that homosexuals must be compared with a group of nonhomosexuals and assessed in terms of the relevant features—normal organ function according to the BST—which implies homosexuality should not be accepted as a reference class. Any XST-variant utilizing a contentious condition as a specification of reference classes can be discounted this way.

Now, this strategy will only work if it does not generalize to the reference classes that Boorse and others do employ: species, sex, and age. Kingma asked for a justification for rejecting theories of health with *more* reference classes; she could also have demanded a justification for rejecting theories with *less*. It is obvious that eliminating any of these three specifications (species, sex, and age) has implications that violate a range of paradigmatic cases of healthiness and pathology, which is a good enough reason to reject the suggestion. But still, the worry might persist that the argument about candidature for reference classes equally disallows the reference classes that Boorse and others do accept.

This last worry is unfounded. If we wanted to know whether it is healthy or pathological to be a member of a certain species, we would indeed require a reference class that is not specified by species, just like finding out whether myopia is a pathology requires a reference class not limited to people with myopia. But, it seems very odd to ask whether it is healthy to be a human being, or a parrot, or a dandelion—does it make sense to ask whether it is healthy or pathological to be a member of these species? It seems clear that health and pathology are species-specific concepts: health and pathology correspond to ways of being *for* a human being, *for* a parrot, *for* a dandelion, and so on. The same goes for sex: the question whether it is healthy or unhealthy to be a man or a woman, a rooster or a hen, a boar or a sow, is a strange question, to say the least. There is no need for a reference class without specification of sex in order to determine the health-status of sexes, for it is not sensible to ask whether "being male" or "being female" as such is healthy or pathological. In the case of age, I admit there does seem to be a sense in which it is meaningful to ask whether it is healthy or unhealthy to be of a certain age, probably because we generally tend to associate good health with youth. Also, in the case of age, however, we do not typically inquire whether "being in one's early 40s" *is* healthy or

pathological. We ask whether *for* a (fe)male human being in one's early 40s one is healthy, not *whether* being in one's early 40s *is* unhealthy. Because questioning the health status of properties like species, sex, and age is at odds with these basic intuitions about health, I suggest the argument used above does not generalize to the properties Boorse and others do accept as reference classes.

VII. CONCLUSION

Boorse's BST of health aims to be a naturalistic theory of health, meaning it rejects the view that personal or cultural values determine whether one is healthy or not. Kingma's objection that use of reference classes renders the BST (and other naturalistic theories using reference classes) circular and value-laden has been met with a threefold response.

First, I argued that the BST is not circular by distinguishing the meaning of terms in a given definition from the methods or procedures of justifying these definitions. While the meaning of terms is not allowed to require a comprehension of the phenomena they are supposed to elucidate on pains of circularity, I pointed out that justificatory practices may utilize prior beliefs about cases of health and pathology without thereby rendering the theory circular.

Second, I argued against the normativist standpoint that the initial sorting between healthy and pathological conditions occurs on the basis of subjective and cultural values. The latter are variable in the way that the distinction between health and pathology is not; pathologies can be positively valued and healthy conditions can be negatively valued. Although evaluative attitudes may roughly correspond to the distinction between health and pathology and provide a first form of access to the phenomena, there are good reasons to think that the reality of these phenomena obtains independently of these attitudes. Naturalists may therefore appeal to the objective nature of health and pathology to justify their choice of reference classes.

Third, the claim that values play a role in favoring one theory over another has been met with a double rejoinder. As in all other areas of philosophy, choice between rival theories need not take place on the basis of evaluative commitments: factors like paradigm cases of the phenomenon in question, theoretical simplicity and uniformity, explanatory power and specificity, give enough reason to favor one theory over another. In addition, I argued that we also have a fairly strong intuitive grip on which properties can serve as a reference class and which cannot: when we want to question the health status of some condition (as in all of Kingma's examples), this condition should not be accepted as a specification of reference classes, precisely because that makes it impossible to assess the health status of that condition.

I think plenty of problems remain for Boorse's BST and that, all things considered, there are better naturalistic theories of health available. I conclude that the objections formulated in Kingma (2007) against Boorse's proclaimed

naturalism—which apply not just to Boorse's BST but also to other theories of health that equally employ reference classes and purport to be naturalistic—are unfounded.

NOTES

1. One example is Werkhoven's (2018) dispositional theory of health. On this view, an organism's health consists in the ratio of what an organism can do compared to what it could maximally do, given its reference class (also constituted by species, sex, and age).

2. Boorse makes a sharp, but much overlooked, distinction between "theoretical health" and "practical health": the former aims to capture the value-free scientific use of the concept; the latter is based on factors like treatability of pathologies, and evaluative judgments about which conditions are undesirable. See Boorse (1977 542, 1997, 45–53, 2014, 23).

3. The BST is normative in another sense: it measures a way of functioning against a norm—the norm being statistical normality. The type of normativity that Boorse denies, and which Kingma accuses him of being committed to after all, is a distinctly *subjective* type of normativity. The accusation is that *personal* and *cultural* norms play a decisive role in judgments about health and pathology. Normativists like Engelhardt (1976, 1986) Agich (1983), Cooper (2002), Goosens (1980), Nordenfelt (1993, 1995, 2001), and Reznek (1987) typically endorse this latter view: the idea that subjective and cultural values are constitutive of health. Denying this form of normativity does not necessarily exclude another form of objective, mind-independent normativity. This point has also been made by Simon (2010) and Broadbent (2017). The way in which the BST does remain normative in my view is what Broadbent calls, derived from Stempsey (2000), "Value Dependent Realism."

4. I also draw from Kingma (2013, 2014) in this portrayal of her position.

5. This is a paraphrase of Kingma's own definition of normativism: "the view that health and disease are primarily value-*laden* concepts, representing or expressing what we value and disvalue, or how we think people ought—morally and/or socially—to be" (2014, 591).

6. She writes: "Boorse's reference classes may result in a naturalistic account of health and disease, that is, a value-free description of health and disease, but this does not mean that those concepts are not social in origin or implicitly reflect norms and values" (Kingma, 2013, 51).

7. I take it that this type of response amounts to what Kingma calls the "more modest" claim that Boorse must make. She writes: "Boorse must adopt a more modest claim: once the reference classes are fixed the BST gives an accurate and value-free analysis of health and disease. In other words, once reference classes are fixed the BST does not appeal to social judgements to move from the facts about a case to a judgement about its health status" (Kingma, 2007, 132). Surely though, even if Boorse were to make this concession, the BST would remain a value-free naturalistic theory of health. I should add that I disagree with Kingma's suggestion here on all fronts: there are good reasons to think that reference classes are not fixed on the basis of subjective values, and that the BST is not an accurate analysis of health and disease.

8. See Boorse (1977, 544–545).

9. Lung cancer is a disease whenever or wherever it occurs, whether it be in a modern European woman, a male Inuit, or a prehistoric caveman—it does not matter whether anyone ever negatively valued the condition.

10. These examples are mentioned by Boorse (2011, 21).

11. As Boorse (2014, 692–693) helpfully points out, this way of phrasing the objection echoes a point made by DeVito (2000). DeVito also argues that "when there are multiple conceptions of the same thing, we can choose one of these concepts as the 'correct' or accepted one. Values can enter at this choice level" (2000, 541–42).

12. It does help, of course, by showing that there is a fact of the matter about which theory is correct.

ACKNOWLEDGMENTS

I am very thankful to Martin Lipman and Jesse Mulder for their helpful comments on earlier versions of this paper.

REFERENCES

Agich, G. J. 1983. Disease and value: A rejection of the value-neutrality thesis. *Theoretical Medicine* 4(1):27–41.

Boorse, C. 1975. On the distinction between disease and illness. *Philosophy and Public Affairs* 5(1):49–68.

———. 1976. What a theory of mental health should be. *Journal for the Theory of Social Behavior* 6(1):61–84.

———. 1977. Health as a theoretical concept. *Philosophy of Science* 44(4):542–73.

———. 1997. A rebuttal on health. In *What Is Disease?* eds. J. M. Humber and R. F. Almeder, 1–134. Totowa, NJ: Humana Press.

———. 2011. Concepts of health and disease. In *Handbook of the Philosophy of Science, vol. 16 Philosophy of Medicine*, ed. F. Gifford, 13–64. New York: Elsevier.

———. 2014. A second rebuttal on health. *Journal of Medicine and Philosophy* 39(6):683–724.

Broadbent, A. B. 2017. Health as a secondary property. *British Journal for the Philosophy of Science* 70(4):927–52.

Cooper, R. 2002. Disease. *Studies in History and Philosophy of Biological and Biomedical Sciences* 33(2):263–82.

DeVito, S. 2000. On the value-neutrality of the concepts of health and disease: Unto the breach again. *Journal of Medicine and Philosophy* 25(5):539–67.

Engelhardt, H. T. 1976. Ideology and etiology. *Journal of Medicine and Philosophy* 1(3):256–68.

———. 1986. *The Foundations of Bioethics*. New York: Oxford University Press.

Goosens, W. 1980. Values, health and medicine. *Philosophy of Science* 47(1):100–15.

Kingma, E. 2007. What is it to be healthy? *Analysis* 67(294):128–33.

———. 2013. Health and disease: Social constructivism as a combination of naturalism and normativism. In *Health, Illness and Disease: Philosophical Essays*, eds. H. Carel and R. Cooper, 37–56. Durham, United Kingdom: Acumen Publishing.

———. 2014. Naturalism about health and disease: Adding nuance for progress. *Journal of Medicine and Philosophy* 39(6):590–608.

Nordenfelt, L. 1993. *Quality of Life, Health and Happiness*. Aldershot, United Kingdom: Ashgate Publishing.

———. 1995. *On the Nature of Health: An Action-Theoretic Approach*. 2nd rev. ed. Dordrecht, The Netherlands: Reidel.

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

Reznek, L. 1987. *The Nature of Disease*. London, United Kingdom: Routledge and Kegan Paul.

Simon, J. 2010. Beyond naturalism and normativism: Reconceiving the 'disease' debate. *Philosophical Papers* 36(3):343–370.

Stempsey, W. E. 2000. *Disease and Diagnosis: Value-dependent Realism*. Dordrecht, The Netherlands: Springer.

Werkhoven, S. 2018. A dispositional theory of health. *The British Journal for the Philosophy of Science* [On-line]. Available: https://doi.org/10.1093/bjps/axy003 (accessed October 21, 2019).