The Journal of Medicine and Philosophy, **47**: 530–539, 2022
https://doi.org/10.1093/jmp/jhac012
Original Research

OXFORD

A Critique and Refinement of the Wakefieldian Concept of Disorder: An Improvement of the Harmful Dysfunction Analysis

Emmanuel Smith*
Florida State University, Tallahassee, Florida, USA

*Address correspondence to: Emmanuel Smith, PhD, Philosophy Department, Florida State University, Diffenbaugh 001A, Tallahassee, FL 32306, USA. E-mail: jes18h@fsu.edu

One way in which bioethicists can benefit the medical community is by clarifying the concept of disorder. Since insurance companies refer to the DSM for whether a patient should receive assistance, one must consider the consequences of one's concept of disorder for who should be provided with care. I offer a refinement of Jerome Wakefield's hybrid concept of disorder, the harmful dysfunction analysis. I criticize both the factual component and the value component of Wakefield's account and suggest how they might be improved. I propose that the factual component should be statistical variation analyzed in multilevel and chronological compilations of physiological data. I propose that the value component should prioritize the individual's authority regarding the experience of suffering from a physiological condition. My account preserves the insight of using a factual and a value component while avoiding the problems that Wakefield's original account faces.

KEYWORDS: *diagnosis, disorder, normal/abnormal, suffering*

I. INTRODUCTION

If I have a mild headache, do I have a disorder? No? What if I have had this headache for the past 15 years? What if I once had the suspicion that someone was out to kill me? Does that mean I have a mental disorder? But, what if I often hallucinate someone threatening to kill me? Somewhere in between the extremes in these examples lies the demarcation for what should count as a disorder and what should not. Identifying this demarcation is important not only for the purpose of accurate diagnoses but also for providing guidance to both medical and insurance personnel in knowing how best to help those who should receive care. In fact, as Jerome Wakefield and Michael First (2013) have noted, most disputes over accurate diagnoses revolve around this distinction between disorder and normality. My purpose in writing this paper is to propose a modified version of Wakefield's two-component account of the concept of disorder, the two components being a factual component and a value component that together define disorder.[1] Although I am largely in agreement with Wakefield's proposal, I argue that it can be strengthened. I first describe and critique both components of Wakefield's account of disorder. Next, I propose refinements to remedy the ailments of Wakefield's account, thereby presenting a modified account of disorder, which is followed by responses to potential objections. I conclude by discussing some implications of this account of disorder for psychiatry in particular.

II. WAKEFIELD'S FACTUAL COMPONENT

Wakefield proposes a "middle" position between value-laden-only accounts of disorder on the one hand, and value-free accounts on the other. He attempts to do so by providing two components, one of which is value-laden and the other of which is factual. The reason for using both a value-laden criterion and a value-free criterion is to preserve the important ideas that the two ends of the spectrum capture, while at the same time avoiding the problems each of the theories face. For example, Wakefield believes that the value account captures the importance of social concern and involvement regarding disorder, whereas the factual account captures the importance and proven effectiveness of scientific research. Nevertheless, he believes both accounts to be inadequate: the value account places the definition of disorder into the hands of a capricious culture (neglecting scientific knowledge altogether), while all of the factual accounts offered prior to Wakefield's proposal depended on inadequate criteria, such as statistical deviance or failure to promote longevity and reproduction.[2] Thus, instead of relying solely on either a factual or a value component, a satisfactory account of disorder should incorporate both. As Wakefield put it,

> The concept of disorder must include a factual component so that disorders can be distinguished from a myriad of other disvalued conditions. On the other hand, facts alone are not enough; disorder requires harm, which involves values. Thus both values and facts are involved in the concept of disorder. (1992, 381)

So, what are the factual and value components that Wakefield utilizes? In short, he presents his factual component in terms of *dysfunction*, by which he means the failure of an organ to function in the way that it was designed to function as a result of natural selection. On the other hand, Wakefield bases his value component on a notion of *harm*, defined in terms of whatever a society deems undesirable (Wakefield, 1992, 373).

Consider, then, Wakefield's factual component. Many of the factual accounts of disorder that preceded Wakefield utilized one notion or another of dysfunction. These were found unsatisfactory for various reasons. In order to improve on these accounts, Wakefield (1992, 383) suggests that one understand the natural function of an organ to be whatever function the organ was naturally selected to perform. Dysfunction, then, would be the failure of an organ to perform its naturally selected function.

To explain how evolutionary theory is supposed to elucidate an organ's natural function, Wakefield considers a comparison between the naturally selected function of an organ and the human-selected function of an artifact. Of course, from one perspective, the function of, say, a watch can be explained by a description of the structure and interactions of all its components: the gears, screws, etc. In Aristotelian terms, this would be from the perspective of the material and formal causes. Now another way to explain the function of a watch is to appeal to the plan of the artisan who designed and built the watch. One can ask, "What is this watch supposed to do? What is its purpose?" to which the reply is, "This watch was designed to display the time accurately." Again, in Aristotelian terms, this would be the perspective of the final cause. Focusing on this latter perspective, Wakefield notes that, with regard to artifacts, the order of explanation can work both ways. That is, the design plan in the mind of the artisan can explain what the watch was built to do, while the function of the finished product can illuminate why the watch was designed and built.

Similarly, Wakefield contends, observing how human livers function today can explain why they evolved with the particular structures and mechanisms they manifest. In other words, just as the functioning of a watch can indicate that it was made to display time accurately, so also can the functioning of organs or biological mechanisms explain their existence as a result of evolution through natural selection. It is this explanatory connection on which Wakefield bases his conception of natural function: "A natural function of an organ or other mechanism is an effect of the organ or mechanism that enters into an explanation of the existence, structure, or activity of the organ or mechanism" (1992, 382).

Applying this definition of natural function, it becomes a matter of empirical research into the natural functions of organs and mechanisms that allows one to identify dysfunction. For example, if the

result of the evolutionary process (thus far) is that a liver stores and releases glycogen (among other things), then that is one of the liver's natural functions and a liver that does not do so is dysfunctional. Thus, utilizing this concept of natural function and dysfunction, Wakefield believes he has the factual component of his definition of disorder.

Given that Wakefield's theory rests on a division of labor between a factual component and a value component in defining disorder, an obvious objection would be to argue that the two components are not actually distinct; that is, one might argue that Wakefield's factual component is inescapably normative. For instance, Scott DeVito (2000) has argued that Wakefield's factual component is not value-free, since he makes an evaluation at the level of his choice of a criterion for dysfunction. In other words, Wakefield's factual component is value-laden in that it is an evaluative decision to choose natural function over something else like quality of life or lack of pain. Now, one might think that based on this understanding of value according to which simply choosing one option over another counts as introducing value into a theory is trivial. That is, one might object that based on this understanding of value it is impossible to avoid a value-laden theory, but that this is inconsequential for a theory. DeVito might agree, however, and note that Wakefield's introduction of value into his factual component is particularly problematic because it is a choice concerning what is or is not *dysfunctional*. DeVito might claim that something is dysfunctional if it does not function in the way in which it *should*. In other words, dysfunction has a normative aspect that introduces value into a criterion that is supposed to be value-free. It may indeed be inconsequential to introduce value into a theory by choosing among, say, different orders in which to solve calculations. It is much more significant for Wakefield to introduce value into his factual component by choosing a criterion with which to identify whether an organ or biological mechanism is functioning as it should.

"So what?" you might say. "Perhaps Wakefield could grant that his factual component is not value-free in any important sense after all and just present his reasons for why his choice is the best among the various options." In fact, Wakefield need not even grant that the type of value introduced at the level of selection of a component means that the component is no longer factual. Wakefield could maintain that his factual component is a descriptive account, not a prescriptive one, as DeVito charged. For, rather than holding that a biological structure, process, or mechanism was naturally selected for the purpose of functioning in this or that way, one could instead hold that the phrase "natural function" is only meant to pick out whichever structure, process, or mechanism persisted through time via the process of natural selection.

However, I find two other worries to be problematic for Wakefield's theory. The first criticism of Wakefield's factual component is aimed at demonstrating that his idea of a naturally designed function is not by itself sufficient to identify dysfunction. The reason is because his component is too narrow insofar as it defines dysfunction by appeal *only* to natural selection through random genetic mutation. Yet, evolutionary theory alone does not provide an adequate basis by which to identify dysfunction. For, according to the predominant interpretation of evolutionary theory, natural selection through genetic mutation is a *random process*. As such, whichever genetic mutations occur do so randomly, which is to say that such mutations were not designed or planned and did not occur *in order to achieve some purpose*.[3] So, Wakefield's understanding of natural function *cannot* suppose that organs and biological mechanisms evolved *in order* to perform a particular function.

Yet, recall that Wakefield defines natural function as an effect that enters into the explanation of the existence, structure, or function of an organ or biological mechanism. The term "explanation" in this definition can be understood in two ways, neither of which is favorable to Wakefield's proposal. As just mentioned, "explanation" could be understood in terms of an effect that explains *why* some organ or biological mechanism evolved—that is, the explanation could be understood in terms of purpose. Now according to evolutionary theory, there is no "why" with regard to the existence, structure, and function of such things; whatever evolved did so randomly—for no purpose. On the other hand, "explanation" could be understood in terms of how an effect enabled an organism to survive through the process of natural selection. However, this sort of explanation implies nothing with regard to how an organ or biological mechanism *should* function, since there is a variety of ways in which such things could have evolved that would have allowed the organ, biological mechanism, or organism to survive.

Note that each, say, human heart is to some extent unique, with a large degree of similarity between one heart and another. Yet, if there is a small percentage of human hearts that do not share much

similarity with other human hearts due to certain genetic mutations, then why not say that those are the hearts that are functioning naturally, while the rest of the hearts are dysfunctional? Furthermore, if the small percentage of hearts were just as evolutionarily fit as the majority, then there would be no justification to which Wakefield could appeal in order to adjudicate which hearts should be considered dysfunctional and which not dysfunctional, since, according to evolutionary theory, biological structures, mechanisms, etc. do not evolve according to some plan or for some purpose.

Hence, the second criticism of Wakefield's factual component: he implicitly relies on statistical variation, which provides the foundation for his idea of natural function. In other words, in order to define the boundaries at which lie the distinction between dysfunction and natural function among the various ways in which biological structures and mechanisms have evolved, Wakefield must utilize a notion of statistical deviation from a norm, which, although he acknowledges (2011, 15) that it could be a useful *guide* for identifying natural functions, is a tactic he explicitly rejects (1992, 377–78). To be clear, my claim is not that Wakefield errs in utilizing a notion of natural function. There is nothing problematic for his factual component in attending to a biological mechanism, structure, etc., that has persisted in a species as a result of happening to confer an advantage for survival. My point is that such a notion of natural function is inextricably associated with biostatistical variation/deviation.

One way in which Wakefield might respond to these objections would be to say that I have misunderstood his idea of a natural function. That is, he could agree that random genetic mutations have no purpose for which they occur but could maintain that a natural function enters the picture only when such genetic mutations have benefitted a species such that they persist and become widespread within that species. In this way, a genetic mutation, or several genetic mutations, become involved in a natural function in an organism only by spreading throughout the species through promoting survival.

However, the only way to identify which genetic mutations (and, at higher levels, which biological functions, organelles, morphological structures, etc.) acquire the status of natural function in this way is to attend to their biostatistical predominance.[4] In other words, in order to determine if some biological mechanism, process, structure, etc., is a natural function for a species, it is necessary to establish that it is statistically common for that species. Similarly, in order to determine if some biological mechanism, structure, etc., is functioning contrary to its natural function, one must know how the majority of similar mechanisms, structures, etc., typically function (but, as both Wakefield and I acknowledge, this identification alone is not sufficient for a satisfactory concept of disorder: there must also be a value component).

Yet, Wakefield most certainly would deny that his idea of natural function relies on deviation from a biostatistical norm. Indeed, he might hold that such a reliance would only introduce problems for the proposal. For example, one could argue that simply identifying biostatistical abnormalities does nothing to elucidate natural function, since there are plenty of examples in nature of phenomena that are biostatistically abnormal but in which there are clearly no dysfunctions, such as the disproportionate ratio of female to male bees in a colony. On the other hand, there are plenty of examples of phenomena that are biostatistically normal but for which we would not want to deny the presence of a dysfunction, such as the percentage of individuals in a population who have had a cavity or who have experienced a fever.

In response, I agree that if Wakefield's theory rested on only a notion of biostatistical deviation, then it would lead to counterintuitive results. This is why there must also be a value component that informs the application of the factual component as well as provides a necessary criterion for the attribution of disorder. However, not only do I maintain that Wakefield's factual component incorporates biostatistical variation, but, as I argue later in this paper, I also contend that this incorporation is for the best. Briefly, though, the typical examples presented against the use of biostatistical deviation miss the mark: it is true that female bees are statistically rare compared to male bees in a colony and that this is not a disordered state of affairs, but we know that this is not a disordered state of affairs because most, if not all, bee colonies have such a ration of female to male bees (hence the use of biostatistical normality and deviation through a comparison of the *relevant* data sets).[5] Similarly, conditions such as having a fever or having a cavity, when examined at the cellular level, are properly labeled dysfunctional, and we know this (as I argue later in this paper) through the comparison of *relevant* data sets.[6] So, my conclusion at this point is that the factual component of Wakefield's theory of disorder is in need of refinement.

III. WAKEFIELD'S VALUE COMPONENT

An organ or biological mechanism can be dysfunctional, according to Wakefield's theory, without qualifying as a disorder. For example, color blindness would be considered dysfunctional (since cone photoreceptor cells are thought to have evolved to respond to stimulation from light waves of certain frequencies), but it would be wrong to call people with this genetic mutation disordered. Such people are able to live within any society just as well as people who are not color-blind, and not being able to perceive red and green is *usually* not painful, inhibiting, etc. This is why Wakefield utilizes a value component in addition to his factual component. It is not enough to say that whatever is dysfunctional is disordered, since that would result in such a broad category that it would include many things that, intuitively, are not disorders. Instead, one must introduce a value component that will allow one to identify which dysfunctions are disorders and which are not. As Wakefield puts it, "To be considered a disorder, the dysfunction must also cause significant harm to the person under present environmental circumstances and according to present cultural standards" (1992, 383). That is, a dysfunction is a disorder only if it is harmful to the individual, and something is harmful to an individual only if it is considered to be harmful by the society in which that individual lives (Wakefield, 2017, 314).

An example of a disorder, then, would be ALS (i.e., amyotrophic lateral sclerosis), since it involves both dysfunction of the nervous system, and society would consider ALS to be harmful. As Wakefield acknowledges, though, any account of disorder will encounter difficulties in yielding judgments for certain cases. For example, homosexuality might be considered a disorder in some societies (e.g., a predominantly conservative Christian society) but not in others (e.g., a predominantly secular society), since a society that considers homosexuality harmful might be able to tell an evolutionary story according to which homosexual behavior is not a natural function.[7] Nevertheless, Wakefield maintains that, despite the inevitability of some vague cases, his value component is the best choice both for informing which dysfunctions should be considered disorders, as well as for involving a society in the care of individuals with disorders.

However, the example regarding homosexuality is a symptom of a larger problem. If something can be considered a disorder only if some society considers a dysfunction harmful, then what is and is not a disorder is largely (though not entirely) dependent on the majority opinion of a particular society. Yet, societies are notoriously unstable, changing their predominant values often just as quickly as there are changes in leadership. Also, in recent times, societies are so cosmopolitan that society-wide consensus is difficult to achieve (what an understatement!).

The problem, in short, is that if the concept of disorder is largely dependent on whatever a society deems harmful, then the concept of disorder could be susceptible to change just as quickly (and as arbitrarily) as cultures change. This means that something that is considered to be a disorder today might not be considered a disorder 10 or 20 years from now. Consider again the example of ALS: if the concept of disorder was dependent on whatever a society thinks is harmful, assuming for the sake of argument that one could make sense of the mishmash of differing opinions within a society about what is and is not harmful, then it is possible (though unlikely) that a society could change its values such that ALS is no longer considered harmful and thus not a disorder. This would be the case even if human physiology and environmental factors remained constant. However, it is obvious that those who have ALS are disordered. Although a theory might be excused for having difficulty in accounting for controversial cases, a theory that allows for the possibility of counterintuitive judgments for the clear cases is in need of revision.

Wakefield might respond to this criticism by contending that such radical changes in the opinions of societies are so improbable as not to matter when considering the merits of his theory.[8] He could concede that cultural standards are likely to change a small amount over time, resulting in a few small cases in which some condition goes from being considered disordered to healthy or vice versa. However, he could hold that it is plausible to assume that societies are made up (mostly) of rational agents who make informed decisions on matters regarding health and medicine; and, if one can assume that societies make rational decisions, then one can claim that the opinions of societies concerning disorders will be stable over time. For support, Wakefield could note that there are certain clear examples of conditions that societies have unanimously considered to be disorders throughout time, such as severe psychosis (even if it has not always been called by the same name or understood

in the same way) or ALS. Moreover, he might note that the reason why it is so intuitive that something like ALS is harmful is because this disorder involves harm that is so basic and universally accepted that it would be unreasonable to question its status as harmful. Indeed, Wakefield could contend that regarding the vast majority of conditions, when it comes to the concept of harm narrowly confined to the concept of disorder, to suggest that social mores are liable to change is simply pointless.

In response, note first that even if it were true that societies are likely to remain stable over time regarding their concepts of harm, it would still be the case that the basis for determining whoever is and is not disordered (specifically, whoever experiences genuine harm) is the majority opinion of a society, not the individual who claims to be experiencing harm. However, since societies are heterogeneous, a society will have minorities who disagree to some extent with the majority opinion. This implies that if there is a group that holds a minority opinion in a society and believes that condition X is harmful, and hence a disorder, whereas the group that holds the majority opinion does not believe condition X is harmful, and hence not a disorder, assuming that both groups agree that condition X is dysfunctional, then the only thing that the majority can appeal to in order to support this difference is *the way they think about condition X*. This would result in an impasse where those whom the minority consider to be disordered are not treated as disordered by and do not receive care from the majority.[9] It is somewhat analogous to when a child asks her parent why punching her sibling is wrong and the parent responds by saying "Because I said so!"

Another reason why it is problematic to place more power concerning legitimate attributions of harm in the hands of the society rather than in the hands of the individual who claims to experience harm is because, *in most cases*, it is the individual who has the most authoritative perspective on the harm. Of course, there are some cases, such as infants and the severely mentally disabled (among others), where the individual in question is incapable of clearly communicating the experience of harm or the lack thereof. This is why the proverbial "pendulum" should not be swung in the complete opposite direction such that societal input concerning harm is disregarded. Nevertheless, just as, in most cases, the individual has the most authoritative perspective on, say, which thoughts he/she is thinking, so also, in most cases, does the individual have the most authoritative perspective on the presence or absence of the experience of harm.

Finally, one need not grant that societal values are mostly stable over time. Just consider how much value judgments in the United States have changed over the span of its short existence. For example, around the time of the founding of the nation, individuals with autism were considered disordered, whereas recently there have been pushes to treat those with autism as neurodiverse (i.e., not disordered).[10] It may be that those who are leading this push to treat autists as neurodiverse are correct in their position. However, the arguments that these people (e.g., Dr. Temple Grandin) advance highlight the importance of placing more emphasis on the perspective of the individual when it comes to suffering. Note, however, that in drawing attention to the lack of stability for general societal values I am not thereby also making the claim that all societal value judgments are equally unstable. Instead, I discuss the lack of stability for general societal values for two reasons: first, in order to motivate the possibility (however unlikely) of significant changes in societal judgments concerning harm, and second, in order to note that even relatively stable societal judgments (such as judgments about harm) can change "at the fringes" and, when a society is the *only* authority regarding the legitimate attribution of harm, such changes "at the fringes" could affect many.

One must recognize that no society is stable or unbiased enough with regard to majority values to ensure that the scope of harm for Wakefield's concept of disorder would encompass *everyone* who deserves consideration. If his value component could be implemented in an ideal global society, then perhaps relying so heavily upon cultural opinion for the concept of disorder would not be problematic; however, no society is ideal, and each is too whimsical to entrust with so much power (i.e., power that can be abused) with which to define the concept of disorder. Indeed, Wakefield (2014a, 353) has recently noted how a society may abuse its power over the concept of harm, such as when a society defines harm in relation to one's failure to meet unreasonable, socially constructed ideals. In such cases, Wakefield is correct to allege illegitimate attributions of disorder due to the lack of dysfunctions (i.e., failure to meet his factual criterion). Nevertheless, I maintain that an appropriate counterbalance to such potential abuses is not only the required satisfaction of a factual component to the concept of

disorder, but also privileging the perspective of the individual rather than the perspective of a society when it comes to the experience of suffering. In summary, I conclude that neither Wakefield's factual component nor his value component is satisfactory. In the next section, I offer refinements of these components that I believe provide improvements for Wakefield's concept of disorder.

IV. A NEO-WAKEFIELDIAN ACCOUNT OF DISORDER

Briefly, the refinement for the factual component involves statistical deviation identified by analyzing both multilevel and chronological data sets. The refinement for the value component is based on placing more emphasis on the individual's authoritative perspective when it comes to the experience of suffering. However, these two components are not intended to be applied independently; instead, each component is used to inform the application of the other. Whenever some condition meets both criteria, it qualifies as a disorder. Furthermore, it is important to understand that where the criteria are not both met, but instead there is either a biostatistical deviation from the relevant norm that does not involve suffering or there is suffering that does not involve a biostatistical deviation from the relevant norm, a failure to meet both criteria does not nullify the satisfaction of one or the other criterion. Also, satisfying only one or the other of the two criteria does not mean that the individual under consideration would not benefit from treatment, such as talk therapy or music therapy. (However, I exclude pharmaceutical treatments in most cases for such people.)

Let us focus first on the factual component. As argued above, Wakefield's original factual component implicitly utilizes a notion of biostatistical variation in order to identify the natural function of an organ, biological mechanism, etc. That is, since organisms do not evolve *identically*, but instead evolve in *roughly similar* ways, in order to have a useful concept of natural function, one must confine the range of roughly similar organs and mechanisms so as to exclude those organs and mechanisms that have a certain amount of dissimilarity when compared to other samples. Furthermore, even when the focus is shifted to the causal or evolutionary history of a certain mechanism, process, structure, etc.,[11] one must still rely on biostatistical data that identifies which mechanisms, processes, structures, etc., proliferated throughout the species by the process of natural selection.

However, as Wakefield notes, there are problems with relying on statistical variation to inform one's concept of disorder. For example, there are several conditions that are statistically abnormal but clearly are not disorders, such as high IQ or very strong muscles. Additionally, some conditions are so prevalent in some areas that they are statistically normal even though they clearly are disorders, such as, perhaps, obesity in some American cities (Wakefield, 1992, 377). In other words, one can ask, "Statistically normal/abnormal with regard to what?" Osteoporosis is a normal condition if one is considering only a narrow set of statistics regarding 90-year-old women. Similarly, a 6′8″ Dutch man is an extreme outlier when compared to the average height of an Indonesian.

In response to this latter problem of justifying the scope of the data that one considers for statistical analysis, I propose taking into account data across a variety of parameters, from the microphysical to the organismic and species scales, as well as across time.[12] More specifically, the scientific community accrues data regarding the structures and functions across time of organelles, cells, blood vessels, organs, organ systems, organisms, etc., all of which form Gaussian curves specific to each data set. These data sets can then be analyzed in isolation as well as in comparison to similar data sets and to data sets at earlier or later stages of development (and doing this using a principle of comparing sets that are *relevantly* similar to one another).[13]

Thus, with such a large amount of data, one will have both an appropriately broad and an appropriately narrow perspective that are sufficient to account for every type of condition.[14] For example, the relevantly similar data set for analyzing the condition of a 6'8" Dutch man would be other Dutch men of a similar age and born at a similar point in time, since the condition under scrutiny is the height of Dutch men of a specific sort. Likewise, the relevantly similar data set for analyzing a 90-year-old woman with osteoporosis would not be other 90-year-old women; instead, one would perform the analysis at the cellular level, scrutinizing, say, osteocytes and osteoblasts over time, since the condition under scrutiny has to do with the functioning (or lack of functioning) of cells involved in the maintenance and growth of bones. Furthermore, not only should data sets be compared between relevantly

similar individuals, but so also should different data sets be compared between various stages of a single individual's stages of development. (I acknowledge that "big data" for most individuals is idealistic at this point, but the move towards personalized medicine could make this potential beneficial resource a reality.[15])

However, there remain two problems. One is the fact that there are many instances of conditions that are statistically deviant but clearly are not disorders. The second problem is how one might non-arbitrarily set a limit on how far to one extreme or the other a condition can be before it is considered statistically deviant. In response to the second problem, in order for this to count as an objection against using statistical deviation for the factual component, it must include the assumption that the conditions of most human organs, biological mechanisms, etc., are so vague as not to allow for a clear majority and a clear minority. However, for most human organs and biological mechanisms, the number of cases that share a high degree of similarity, while still allowing for each to be slightly unique, far outnumbers the number of cases that are obvious outliers. In other words, it is mistaken to assert that, in relying on statistical analysis of all the relevant data with regard to human beings, one could not clearly identify outliers that deviate starkly from a clearly identifiable majority.[16]

"However," you might say, "even if this is correct, it is still the case that some human beings will exhibit conditions that are both somewhat 'normal' (identified as the majority group) and somewhat 'abnormal' (identified as the outliers), which means that there will be vague conditions for which your factual component cannot account." Before providing an answer to this objection, as well as offering a response to the first problem mentioned above (i.e., how to deal with statistically abnormal conditions that clearly are not disorders), it will be helpful to present the value component of my account, since it has an important role in informing the application of the factual component.

The value component is fulfilled if an individual experiences suffering that is caused directly by a biological condition. I define "suffering" as a physiological state that consists in the subjective experience of unwanted pain, unpleasantness, aversion, or distress.[17] A biological condition may cause an individual to suffer directly if it is the primary and immediate cause of pain, unpleasantness, etc. So, although a biological condition may cause an individual to suffer indirectly if the condition directly causes circumstances external to the individual that are the primary and immediate cause of pain, unpleasantness, etc., such as when an individual experiences physical injuries or depression from social ostracization, both of which can result in a state of suffering, such an experience of externally imposed suffering would not, *on this basis alone*, satisfy the value component. Additionally, the main arbiter who determines whether an individual is experiencing suffering is the individual him- or herself. That is, it is the individual, not the society within which the individual resides, who has the primary authority to determine whether he or she is experiencing suffering. In most cases, the judgment of the individual and the judgment of a society will converge. However, the value criterion must preserve a role for the judgment of a society for those cases in which the individual is incapable of clearly communicating the experience of suffering. Moreover, as argued in my critique of Wakefield's original value component, there must also be a privileged role for the judgment of the competent individual, especially in cases involving psychological suffering.

Now for the objections. First, a condition could be statistically abnormal (e.g., high IQ) and yet not be classified as a disorder if it does not also fulfill the value condition, which is to say that it does not also cause the individual to suffer. For example, an individual may be statistically abnormally short (within the relevant comparison class) but not experience any suffering as a result of such a condition, which means that the person would not be classified as disordered. Second, in response to the objection that the factual component has no way of marking a distinction between disordered and healthy in vague cases (i.e., cases where it is unclear if the condition is statistically normal or abnormal), the value component can serve as the adjudicating factor. That is, regarding those cases that are both somewhat normal and somewhat abnormal, if the condition fulfills the value component, then it can also be considered to fulfill the factual component (i.e., it qualifies as a dysfunction). In other words, if a statistically vague condition causes an individual to experience suffering, then that condition can also be considered dysfunctional based on statistical abnormality.

Is it arbitrary to use suffering as the criterion that determines the boundary between disordered and healthy in statistically vague conditions? No. Note first, though, that most conditions are not

statistically vague, and so there is no need to appeal to the suffering criterion in order to identify which conditions are normal and abnormal. Regarding the charge of arbitrariness, there are two reasons why the suffering criterion is appropriate. First, at least one of the purposes of having an accurate concept of disorder is so that the medical and insurance communities will be prepared to address those who need care, and if any group is an appropriate recipient of care, it is those who experience suffering. Second, since conditions that cause suffering most often are evolutionarily disadvantageous (i.e., it is more likely than not that conditions that do not cause suffering will be naturally selected), and since the majority of a species propagates whatever conditions are naturally selected (since significant genetic mutations are rare), it is legitimate to infer that the majority of a species will most likely have conditions that do not cause suffering. Thus, I conclude that individual suffering is an appropriate criterion to use when determining what is disordered and healthy with regard to statistically vague conditions.

V. IMPLICATIONS

In conclusion, consider two implications that my suggested improvements of Wakefield's concept of disorder have concerning psychiatry in particular. First, certain conditions that are caused by circumstances external to the individual (e.g., depressive mood caused by poverty or loss of a loved one) *and do not result in a clear statistical deviation in some physiological structure or mechanism* would not be classified as disorders, since they would not satisfy the factual component. Thus, in most cases, such individuals should not be treated pharmaceutically for their conditions, although they could benefit from some form of psychotherapy. However, I acknowledge that the medical sciences are explanatorily incomplete at this point in history. Thus, there are cases in which a person reports suffering despite the inability of medical practitioners to identify any abnormal physiological conditions. In these extreme cases, it is reasonable to attempt to alleviate the individual's suffering through the use of pharmaceuticals, if pharmaceuticals are the only effective intervention.

Second, it could be the case that two individuals are similar in that they both exhibit a statistically abnormal condition, but that one individual suffers from the condition, whereas the other individual does not. In this case, although both individuals are physically similar, one would be classified as disordered and the other as healthy. For example, one individual may not experience suffering as a result of her condition and thus not have a disorder (but would still have a dysfunction), whereas a different individual who does experience suffering as a result of his condition would have a disorder. For example, perhaps both individuals have some neurophysiologically identifiable dysfunction, but the former does not suffer, whereas the latter does. Rather than being a problem for my proposal, this implication adds support, since it indicates how the account can accommodate diversity in human experience, potentially resulting in fewer people being neglected by the medical community, as well as fewer people being treated or stigmatized unnecessarily.

The claim I hope to have supported in this paper is that the Wakefieldian concept of disorder is promising but in need of refinement, and that my neo-Wakefieldian proposal provides such refinements. Neither of Wakefield's components in his two-component concept of disorder is *sufficient* to provide medical and insurance personnel with accurate and reliable guidance for diagnoses. The combination of the biostatistical deviation component and the suffering component, as defined in this paper, offers an empirically rigorous and philosophically sound concept of disorder.

REFERENCES

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

Bueno-Gómez, N. 2017. Conceptualizing suffering and pain. *Philosophy, Ethics, and Humanities in Medicine* 12(1):7.

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

Hamburg, M. and F. Collins. 2010. The path to personalized medicine. *The New England Journal of Medicine* 4:301–04.

Hartwell, L., L. Hood, M. Goldberg, A. Reynolds, and L. Silver. 2011. *Genetics: From Genes to Genomes*. 4th ed. New York: McGraw Hill.

Herron, J. and S. Freeman 2014. *Evolutionary Analysis*. 5th ed. Boston: Pearson Publishing.

Mwangi, B., F. Kapczinski, and I. Passos. 2018. Big data and personalized care in bipolar disorder: Recent findings and future perspectives. 20th Annual Conference of the International Society for Bipolar Disorders, 23–24.
Sarto-Jackson, I. 2018. Time for a change: Topical amendments to the medical model of disease. *Biological Theory* 13(1):29–38.
Savarese, R. 2013. From neurodiversity to neurocosmopolitanism: Beyond mere acceptance and inclusion. In *Ethics and Neurodiversity*, eds. C. D. Herrera and A. Perry, 191–205. Newcastle upon Tyne, United Kingdom: Cambridge Scholars Publishing.
Svenaeus, F. 2014. The phenomenology of suffering in medicine and bioethics. *Theoretical Medicine and Bioethics* 35(6):407–20.
Tal, O. 2009. From heritability to probability. *Biology and Philosophy* 24(1):81–105.
Urry, L., M. Cain, S. Wasserman, P. Minorsky, and J. Reece. 2016. *Campbell Biology*. 11th ed. New York: Pearson.
Wakefield, J. 1992. The concept of mental disorder: On the boundary between biological facts and social values. *American Psychologist* 47(3):373–88.
———. 2011. Darwin, functional explanation, and the philosophy of psychiatry. In *Maladapting Minds: Philosophy, Psychiatry, and Evolutionary Theory*, eds. P. R. Andriaens and A. De Block, 143–171. Oxford, United Kingdom: Oxford University Press.
———. 2014a. Social construction, biological design, and mental disorder. *Philosophy, Psychiatry, & Psychology* 21(4):349–55.
———. 2014b. The biostatistical theory versus the harmful dysfunction analysis, Part 1: Is part-dysfunction a sufficient condition for medical disorder? *Journal of Medicine & Philosophy* 39(6):648–82.
———. 2017. Can the harmful dysfunction analysis explain why addiction is a medical disorder? Reply to Marc Lewis. *Neuroethics* 10(2):313–17.
Wakefield, J., and M. First. 2013. Clarifying the boundary between normality and disorder: A fundamental conceptual challenge for psychiatry. *Canadian Journal of Psychiatry* 58(11):603–05.
Wakefield, J. C., and J. A. Conrad. 2019. Does the harm component of the harmful dysfunction analysis need rethinking? Reply to Powell and Scarffe. *Journal of Medical Ethics* 45(9):594–96.

NOTES

1 I use the words "disease" and "disorder" interchangeably in this paper.
2 Cf. Boorse (1975).
3 See Hartwell et al. (2011, 663), Herron and Freeman (2014, 105), and Urry et al. (2016, 469).
4 Cf. the utilization of statistical variation with regard to genetic inheritance in Tal (2009).
5 This example should be confined only to bees that live in colonies and not to the many bee species that are solitary.
6 Cf. Wakefield (2011, 18).
7 Cf. DeVito (2000, 552).
8 Wakefield (Wakefield and Conrad, 2019) has recently made just this sort of claim. That is, he has clarified that cultural standards—not temporary cultural sentiments—are what determines the content of the harm criterion. His clarification, though, does not evade the objections I go on to give in the following paragraphs.
9 The debate over homosexuality in America is a prime example of the opposite situation, namely, the majority of medical professionals think homosexuality is neither harmful nor a dysfunction, whereas a minority of conservative (predominantly Christian) professionals think homosexuality is a disorder (witness: "conversion therapy"). Cf. Wakefield (2014b, 676).
10 See, e.g., Savarese (2013).
11 Cf. Wakefield (2011, 15).
12 Cf. the multilevel and chronological emphases in Sarto-Jackson (2018).
13 Data sets are relevantly similar if most of their categories overlap and they are associated with the physiological state or condition under scrutiny. Categories are things like age, ethnicity, provenance, etc. For example, comparing a data set for ganglion cells in four-year-old Japanese girls with a data set for femur length in 40-year-old Jamaican men would be inappropriate, since these two data sets are irrelevant to one another.
14 See, e.g., Mwangi, Kapczinski, and Passos (2018).
15 Cf. Hamburg and Collins (2010).
16 This, of course, is a matter to be settled empirically. The prevalence of textbooks that teach general facts on human biology, anatomy, physiology, and so on, all of which converge on most things, supports my claim.
17 Similar but more detailed definitions can be found in Svenaeus (2014) and Bueno-Gómez (2017).