

Essentialism Revisited: Evolutionary Theory and the Concept of Mental Disorder

Scott O. Lilienfeld and Lori Marino
Emory University

J. C. Wakefield's (1999) elaboration of his harmful dysfunction analysis (HDA) of mental disorder does little to address previous criticisms (S. O. Lilienfeld & L. Marino, 1995) and instead reveals further conceptual weaknesses in his position. The authors demonstrate that (a) a Roschian analysis can account for the results of all of Wakefield's conceptual experiments and predicts a number of judgments of disorder not predicted by the HDA, (b) the HDA is incapable in many cases of providing a scientifically nonarbitrary distinction between disorder and nondisorder, and (c) the HDA cannot account for failures of cultural exaptations, mismatches between evolutionary design and novel environments, or defenses against threat. The authors argue that the HDA has been convincingly falsified and discuss the failure of essentialistic concepts to resolve controversies in other domains of biological science.

Whenever we have made a word . . . to denote a certain group of phenomena, we are prone to suppose a substantive entity beyond the phenomena. (James, 1890, p. 194)

The long history of attempts to provide an explicit operationalization of disorder, including mental disorder, can be characterized as an ongoing debate between the essentialists and the nominalists. Essentialists believe that natural concepts (i.e., concepts referring to entities in nature, rather than to artifacts) correspond to genuine distinctions in the real world. To them, the demarcation between disorder and nondisorder reflects an underlying "joint" in nature that awaits more precise carving by researchers. Although a number of essentialist criteria for disorder have been proposed, such as structural abnormality (see Kendell, 1986, for a review) and reduced fertility or longevity (Kendell, 1975), they have all been found wanting on logical grounds (Gorenstein, 1984; Wakefield, 1992a). Nominalists, in contrast, believe that natural concepts are mental constructions that bear no direct correspondence to reality and that there are no unambiguous joints in nature to carve. Nominalist criteria for disorder, such as social values regarding a condition's undesirability (Sedgwick, 1982) and statistical deviance (Cohen, 1981) have, however, fared little better than those of the essentialists (Wakefield, 1992a).

In several articles, Wakefield (1992a, 1992b, 1993, 1999) has proposed a largely essentialistic account of disorder,¹ the harmful dysfunction analysis (HDA). According to the HDA, all disorders are harmful dysfunctions, whereby harm is a value judgment regarding a condition's undesirability and dysfunction is the failure of a system to perform as it was designed to perform by natural

selection. Wakefield's definition of dysfunction harkens back to earlier essentialistic accounts of disorder in its emphasis on a natural, in this case evolutionary, demarcation between function and dysfunction. As Wakefield notes, the HDA posits a classical conception in which disorder is operationalized by singly necessary and jointly sufficient features, in this case harm and dysfunction.

We (Lilienfeld & Marino, 1995; see also Lilienfeld, 1995) recently criticized Wakefield's HDA on a variety of grounds and argued that *disorder* is best conceptualized as a Roschian concept (Rosch, 1973; Rosch & Mervis, 1975). Similar accounts of disorder based on Roschian and related family resemblance models have been proposed by several other authors (Dawes, 1994; Kir-mayer, 1994; McNally, 1994). Unlike classical concepts, Roschian concepts lack defining features. Moreover, they possess fuzzy boundaries, so that the question of whether an object or individual belongs to a given category, in this case the category of disorder, is scientifically arbitrary. By scientifically arbitrary, we mean that there is no essentialist criterion or set of criteria (e.g., dysfunction as defined by Wakefield, reduced fertility or longevity; Kendell, 1975) in nature that can be used to definitively distinguish all cases of disorder from all cases of nondisorder. In addition, Roschian concepts are organized around an ideal mental prototype that embodies the central features of the category. According to our Roschian analysis (RA) of disorder, judgments regarding whether persons possess a disorder are based on their similarity to an ideal prototype of a "diseased person."²

The RA, like Wakefield's HDA, represents an attempt to account for consensual judgments of disorder and therefore to answer the question, What do people mean by disorder? Unlike the

Scott O. Lilienfeld, Department of Psychology, Emory University; Lori Marino, Department of Biology and Yerkes Regional Primate Research Center, Emory University.

The authors thank Robert L. Spitzer, Richard J. McNally, and several others for their helpful comments on an earlier draft of this article.

Correspondence concerning this article should be addressed to Scott O. Lilienfeld, Psychology Department, Room 206, Emory University, Atlanta, Georgia 30322. Electronic mail may be sent to scott@ss.emory.edu.

¹ As Wakefield points out, the HDA is not purely essentialistic, because the harm component involves social values.

² Some theorists regard the prototype not as an ideal (i.e., imaginary) individual or object that embodies all of the features of the category but as an actual individual or object that provides the best example of the category (Mervis, 1980). For the sake of simplicity, this distinction will be ignored here.

HDA, however, the RA posits that the consensual meaning of disorder is not underpinned by any single scientific criterion or set of scientific criteria. Consequently, according to the RA, the question of whether certain conditions are disorders or nondisorders has no true scientific answer, although the question of why certain conditions tend to be viewed as disorders or nondisorders is amenable to scientific inquiry.

In the target article featured in this special section of the *Journal of Abnormal Psychology*, Wakefield (1999) criticizes the RA and concludes that the HDA "has no persuasive alternative or even serious rival at this time" (p. 398). In the present article, we demonstrate that Wakefield's criticisms do not withstand close scrutiny and that the RA can accommodate all of his apparent counterexamples. Moreover, we show that the HDA is clearly falsified by Wakefield's own method of conceptual analysis.

The ground rules for our comparison of the HDA and RA are identical to those of Wakefield: We (a) neglect the distinction between disorder and disease (but see Kazdin, 1983 and Lilienfeld, Waldman, & Israel, 1994); (b) focus on the broad concept of disorder rather than mental disorder per se, although we place substantial emphasis on the latter; (c) use Wakefield's (1992b) method of conceptual analysis to compare the validity of the HDA and RA against consensual judgments of disorder; and (d) focus on ontologic, rather than epistemic, questions concerning the validity of the HDA (see Lilienfeld & Marino, 1995). In addition, like Wakefield, we use the term "design" as a shorthand designation to describe the consequences of natural selection but emphasize that our use of this term carries no teleological implications (Lilienfeld & Marino, 1995). Finally, our primary criticisms concern the dysfunction component rather than the harm component of the HDA, as our principal objection to the HDA is the tenet that evolutionary theory can be used to provide a scientific distinction between disorder and nondisorder.

Roschian Concepts: Wakefield's Misunderstandings

Although Wakefield (1999) intends to compare the HDA and RA against consensual judgments of disorder, his analysis of Roschian concepts is plagued by four critical misunderstandings. These misunderstandings undermine his comparisons of the HDA and RA and invalidate the results of many of his conceptual experiments.

First, Wakefield (1999) asserts the following: "They [Lilienfeld and Marino] claim . . . that *disorder* is a mentally constructed category that does not correspond to anything in reality" (p. 397). Moreover, he avers that even if the RA were correct, it would not imply that natural concepts are "mental constructions that do not correspond to any reality" (p. 397). These assertions represent serious misreadings of the RA. In our article, we noted that "the features of Roschian concepts do not arise from thin air: Such concepts emerge largely from repeated experience with real-world entities" (Lilienfeld & Marino, 1995, p. 417). Rosch (1973) similarly emphasized that Roschian concepts often mirror reality to some extent, because they derive largely from encounters with actual stimuli. Moreover, Roschian concepts are organized around prototypes precisely because the distributions of real-world stimuli tend to cluster around certain consistent features (Neisser, 1979).

Wakefield (1999) appears to have confused our claim that the concept of disorder "lacks a clear point of demarcation in the real

world" (Lilienfeld & Marino, 1995, p. 417) with the claim that Roschian concepts bear no relation to reality. For example, most proponents of a Roschian view would contend that the concept of *lake* does not correspond to a genuine point of demarcation in the real world and that distinctions among ponds, lakes, and seas are scientifically arbitrary. Such distinctions are probably made partly on the basis of size and partly on the basis of social conventions guided by particular interests (e.g., whether such bodies of water can accommodate large boats). A Roschian view therefore does not imply that individuals' concepts of *lake* do "not correspond to anything in reality" or are not derived from observations of real bodies of water, because these concepts are substantially shaped by repeated encounters with actual stimuli. Wakefield's error is crucial, as we show later that an RA derived from observations of real-world individuals can account for all of his purported counterexamples.

Second, Wakefield (1999) repeatedly assumes that the features constituting Roschian concepts must be singly sufficient for category membership. For example, several times (e.g., pp. 378, 384, 386, 396), he asserts that a given Roschian feature (e.g., maladaptiveness, inappropriateness) should be sufficient for categorizing a condition as a disorder according to the RA, and then he invokes the failure of this feature to account for shared judgments of disorder to refute the RA. But in fact, an RA does not require that individual features be either necessary or sufficient for category membership (e.g., Cantor, Smith, French, & Mezzich, 1980). Because many of Wakefield's conceptual experiments rest on the assumption that a given Roschian feature must be sufficient to explain disorder, many of his apparent counterexamples are directed at a straw man version of the RA.

Third, Wakefield (1999) consistently assumes that similarity to a Roschian prototype is based exclusively on observable features. Wakefield acknowledges that although "the featural comparison in a Roschian analysis is generally limited to observable properties . . . it is an easy step in defending a Roschian account to include other properties as well" (p. 377). Yet in his comparisons of the HDA and RA, Wakefield repeatedly assumes that similarity of a given condition to a prototype of disorder is based entirely on the extent to which that condition shares superficial features with the prototype. For example, he argues on p. 378 that such conditions as (a) depression and uncomplicated bereavement and (b) dyslexia and illiteracy should be grouped together in an RA because of their observable similarities. Wakefield then uses these comparisons to argue that the RA is incorrect, because only one member of each pair tends to be classified as a disorder. But this conclusion is unwarranted. Recent treatments of family resemblance models emphasize that similarity to a prototype is often based on unobservable features, some of which reflect implicit notions concerning the deeper principles underlying category membership (Komatsu, 1992; Medin, 1989).

Fourth, according to the RA, the classification of a given condition as a disorder is scientifically arbitrary. The RA therefore implies that the question, "Is Condition X a disorder?" is intrinsically unanswerable. Wakefield thus misses the mark when he asserts that our RA implies that "whether a condition is classified as a disorder should be decided on the basis of value judgments" (p. 397). To the contrary, the RA implies that the question of whether a condition is a disorder is scientifically vacuous and should be abandoned, and that societal values are inevitably used

in part to decide *whether this condition merits treatment* (see also Gorenstein, 1984). We say scientifically vacuous because according to the RA there exist no criteria in nature for demarcating disorder from nondisorder. Wakefield finds our "attempt to resuscitate the 'value' view . . . bewildering" (p. 397) only because he confuses our claim that values play a crucial role in who should receive treatment with the incorrect claim that values play an exclusive role in judgments of disorder.

The Reports of My Death Have Been Greatly Exaggerated: The RA and Wakefield's Conceptual Experiments

From where do individuals' Roschian concepts of disorder originate? To begin to address this question, we need to examine the literature regarding the phenomenon of *psychological essentialism* (Medin, 1989). According to Medin and his colleagues, humans possess a tendency to perceive underlying essences in objects and individuals even in their absence. This tendency is especially pronounced with natural (e.g., species) as opposed to artificial (e.g., furniture) concepts, probably because individuals tend to invoke underlying biological attributes (e.g., genetic or physiological factors) for the former categories (Rothbart & Taylor, 1992). As Medin and Ortony (1989) noted, "Psychological essentialism should not be equated with the classical view that concepts are representations of classes of objects that are singly necessary and jointly sufficient for membership" (p. 184). Psychological essentialism often leads individuals to believe that natural concepts are underpinned by entities defined by necessary and sufficient attributes, even though such attributes are nonexistent (Komatsu, 1992; Medin & Ortony, 1989).

Medin and Ortony (1989) contended that natural categories tend to be organized around an *essence placeholder*, which consists of beliefs concerning the underlying properties that all group members share. In many cases, the essence placeholder consists of vague and incomplete suppositions regarding the unobservable essence supposedly possessed by all category members. In addition, it sometimes consists of beliefs "that there are people, experts, who really know what makes the thing what it is, or scholars who are trying to figure out exactly what it is" (p. 185).

In our previous article (Lilienfeld & Marino, 1995), we suggested that Roschian concepts of disorder may be shaped partly by "repeated exposure to conditions perceived as necessitating medical intervention" (p. 417). As Wakefield (1992a, 1999) notes, however, the hypothesis that disorder can be equated with professional concern (e.g., Kendell, 1986; Kräupl Taylor, 1971) is subject to numerous counterexamples. For example, although doctors frequently deliver babies and perform plastic surgery, pregnancy and minor facial defects are not considered disorders. Despite these counterexamples, we suspect that *therapeutic concern* plays a more important role in judgments of disorder than Wakefield implies. Campbell, Scadding, and Roberts (1979), for example, asked groups of physicians, nonphysician academics, and students to listen to a list of conditions (e.g., malaria, schizophrenia) and rate whether each condition is a disease. They found that the most consistent factor predicting disease classification was "the importance of the doctor in diagnosis and treatment" (p. 757).

We propose that individuals' Roschian concepts of disorder derive primarily from both direct (e.g., personal observations) and

indirect (e.g., book and film portrayals) experiences with persons who require or have sought treatment for unwanted medical and psychological conditions and are further perpetuated and shaped by observations of persons who have been labeled as "diseased." From a lifetime of these experiences, we extract a loose set of features shared by such persons, organized around a prototype of the "diseased person." Although Roschian concepts of disorder are derived primarily from experiences with people, they can be extrapolated to explain the perception of diseases in nonhumans (see Wakefield's [1999] Conceptual Experiments 5 and 6).

We further propose that the prototype of the "diseased person" is characterized by an inchoate and ill-formed notion that "there is something wrong with the body or mind that needs to be fixed," as this is the core latent attribute shared by most individuals who seek treatment for medical or psychological problems. This underlying belief, we suggest, functions as an essence placeholder (Medin & Ortony, 1989) that captures our intuitive sense of what diseased individuals have in common. This placeholder is neither necessary nor sufficient for disorder but instead provides an approximate anchor that undergirds the prototype for disorder. The inference that there is something wrong with the body or mind is, we hypothesize, itself based on observations regarding several variables, foremost among which is a sudden or marked decrement in the functioning of physical or mental systems or, in the case of congenital or longstanding conditions, a marked negative deviation from the normally expected functioning of these systems.

According to our RA, the belief that a condition requires treatment should receive considerable weight in judgments of disorder, as this belief is effectively part of the essence placeholder for disorder.³ If individuals perceive that a condition (e.g., intense anxiety) can be easily remedied by means of a nonprofessional intervention (e.g., quitting a stressful job), they will tend not to classify it as a disorder. In contrast, if a condition is perceived as necessitating formal medical or psychological intervention, it is more likely to be classified as a disorder. In addition, the sense that something in the body or mind has changed for the worse should be a moderately to highly weighted feature of disorder, because the perception that something is wrong with the body or mind often derives from a relatively sudden or dramatic decrement in the functioning of one or more physical or mental systems.

Note that our RA is not subject to the same counterexamples as previous accounts (e.g., Kräupl Taylor, 1971) positing professional concern as a defining feature of disorder. For example, our RA would not predict that pregnancy or minor facial defects should be classified as disorders. Pregnancy is not perceived as a condition in which something has gone wrong with the body, or as a condition that needs to be fixed. Minor facial defects are not typically perceived as conditions in which something has gone wrong with the body, primarily because they are not associated with a decrement in functioning. Nor are they generally perceived as conditions that need to be fixed but rather as slight imperfections for

³ We should note that certain conditions are typically perceived as "requiring treatment" even though no effective treatments for them are currently available. Psychopaths, for example, are generally viewed as needing treatment or professional help even though successful treatments for their disorder have yet to be discovered (Lykken, 1995). We are grateful to Robert L. Spitzer for bringing this point to our attention.

which treatment is optional. Because few individuals with minor facial defects seek treatment, such defects possess low “cue validity” (Rosch & Mervis, 1975) for the concept of disorder.

It should also be noted that the intuitive belief that there is something wrong with the body or mind is not simply reducible to dysfunction. Wakefield (1993) proposed that “The idea that something in the person has ‘gone wrong’ is essential to the concept of disorder” (p. 167), a view quite similar to that we have advanced here. Nevertheless, Wakefield (1993) further maintained that this inference is invariably rooted in the failure of a naturally selected system to perform its designed function (i.e., dysfunction). In contrast to Wakefield, we argue that this inference can arise from a number of sources other than a dysfunction of an evolved system. For example, we will later argue that some dysfunctions as defined by Wakefield, namely, those that produce breakdowns of other functions that have not been naturally selected, do not correspond to this intuitive belief. We further argue that some consensual disorders are designed responses to threat. Such disorders, we maintain, do not involve dysfunctions but nonetheless lead to this intuitive belief. In both cases, the belief that something has gone wrong with the body or mind diverges from Wakefield’s concept of dysfunction. Such divergence allows us to compare the validity of the HDA and RA: Do judgments of disorder tend to be more closely associated with dysfunction or with the belief that something has gone wrong with the body or mind?

In his article, Wakefield (1999) presents 12 conceptual experiments that compare the HDA and RA against consensual judgments of disorder. In each case, he maintains that the HDA provides a superior explanation for these judgments, leading him to conclude that the overall pattern of results “disconfirms the Roschian account” (p. 374). We demonstrate below, however, that Wakefield’s dismissal of the RA is premature. Specifically, for each of Wakefield’s conceptual experiments, we briefly explain how the RA can account for consensual judgments of disorder and nondisorder.

Conceptual Experiment 1

Brown eyes are considered to be genetic disorders when due to a spontaneous mutation but not when they have been present all along, because only in the former case is there a perception of a sudden and pronounced decrement in functioning and therefore an inference that something has gone wrong with the organism’s physiology.

Conceptual Experiment 2

Reading disorders are generally considered disorders only when they are due to brain impairment but not when they are due to low motivation, lack of educational opportunity, and so on, because only in the former case is there a perception of a marked or dramatic deviation from expected functioning and thus an unambiguous inference that something is wrong with the brain. Moreover, because individuals do not typically require or seek treatment for low motivation or lack of educational opportunity per se, these conditions possess low cue validity for disorder.

Conceptual Experiments 3 and 8

Individuals infer that there must be something wrong with the physiology of athletes who suddenly lose their physical ability for

no obvious reason but not with the physiology of athletes who experience a decline in ability for obvious external reasons, such as aging or decreased training. Similarly, individuals infer that something has gone wrong with the mental functions of an individual whose IQ has suddenly dropped from 120 to 90 but not with those of an individual whose IQ has always been 90, because only the former individual has experienced a marked and abrupt decrement in functioning.

Conceptual Experiment 4

Currently maladaptive or undesirable traits such as homeliness or being Jewish in Nazi Germany are not perceived as conditions in which there is something wrong with the body or mind, largely because such characteristics are not associated with marked decrements in physical or mental functioning. Moreover, individuals do not typically require treatment for these and other maladaptive attributes.

Conceptual Experiments 5 and 6

Both the moth with a mutation that suddenly results in a change in coloration and the bacterium with a mutated magnetosome that suddenly results in a change in swimming direction tend, unlike their nonmutated counterparts, to be regarded as disordered, because only the former cases result in the inference that something has dramatically and abruptly gone wrong with the organism’s functioning.

Conceptual Experiment 7

The person with a genetic variant resulting in a need for greater amounts of folic acid is not viewed as disordered because individuals perceive that there is no need for treatment or professional intervention. Instead, the damaging effects of this genetic condition can be avoided by a simple alteration in diet.

Conceptual Experiment 9

Chronic abdominal pain is regarded as a disorder only if there is no clear external stimulus, because only then do individuals make the attribution that there is something wrong with the functioning of the body. When there is a clear external stimulus (e.g., a tight belt), individuals perceive that there (a) is no decrement in functioning and (b) is no need for treatment because the pain can be ameliorated simply by eliminating the stimulus. Similarly, when the pain is clearly due to another disorder, individuals conclude that the pain can be relieved by treating the other disorder.

Conceptual Experiment 10

Sneezing is not considered a symptom of disorder if due to dusty basements or other minor irritants because there is no overall decrement in functioning and thus no inference that something has gone wrong with the body. Moreover, individuals perceive that such sneezing does not require treatment and can instead be eliminated by removing the irritants. Sneezing is considered a symptom of a disorder only if it is accompanied by an inference that there is a problem with the body (e.g., flu) that produces a pronounced decrement in functioning.

Conceptual Experiment 11

If Wakefield is correct that fever is no longer considered a disorder, it is probably because it is no longer regarded as a problem that inevitably requires treatment. Indeed, what may have changed is not the perception that fever is a disorder but rather the perception that a fever should always or usually be suppressed by medical means.

Conceptual Experiment 12

Certain conditions (e.g., panic disorder) characterized by inappropriate emotional reactions are considered to be disorders, whereas others (e.g., a tendency to respond with fear to a harmless snakelike object) are not, because only in the former cases is there a clear decrement in functioning or a deviation from expected functioning. Moreover, individuals who exhibit the former, but not the latter, reactions possess a lower threshold than other individuals for reacting to threat and consequently experience more distress and impairment. As a result, they are more likely to require treatment.

Thus, contrary to Wakefield's assertions, the RA can readily account for all of his apparent counterexamples. Nevertheless, because up to this point in the article we have not shown that the RA can account for consensual judgments of disorder that the HDA cannot, we have yet to provide evidence that the RA provides a more conceptually valid (Wakefield, 1992b) explanation for consensual judgments of disorder than the HDA. For example, we have yet to demonstrate that the intuitive belief that there is something wrong with the body or mind includes conditions that are not dysfunctions as defined by Wakefield. We provide evidence for a divergence between this belief and Wakefield's concept of dysfunction later in the article.

There Is No There There: The Lack of Natural Boundaries Problem

A crucial assumption of the HDA is that dysfunction can be validly discriminated from nondysfunction. In slightly different terms, Wakefield (1999) posits that dysfunction is a *taxon*, that is, a nonarbitrary class existing in nature. Taxonicity comes in a variety of forms, including threshold effects, bimodality, bitangentiality, and step functions (Meehl & Golden, 1982). If the distinction between a dysfunctioning and functioning system is nontaxonic, then the *raison d'être* of the HDA is undermined, because the distinction between designed and nondesigned functioning can be made only with reference to nonscientific criteria, such as value judgments. According to the HDA, dysfunction is a scientific concept that is independent of social values (Wakefield, 1992a).

It should be noted that the ability to make a practical distinction between disorder and nondisorder implies nothing about whether this distinction is scientifically nonarbitrary. Wakefield (1999) misunderstands this point when he reviews the physiological bases of essential hypertension, coronary artery disease, and diabetes mellitus, and then asserts that "in none of these cases is there a precise boundary between dysfunction and nondysfunction, yet for practical medical purposes physicians are able to 'adequately distinguish disorder from nondisorder' in all three cases" (p. 379). But the fact that physicians can make a practical distinction between dysfunction and nondysfunction says nothing about the

underlying basis of this distinction. This distinction could be made, for example, on the basis of a scientifically arbitrary demarcation between normality and abnormality (e.g., a standard cut-off for defining high blood pressure) or on the basis of the perceived harmfulness of the condition. In neither case would dysfunction be taxonic, even though a practical decision between dysfunction and nondysfunction can be easily made in most cases.

Wakefield (1999) repeatedly implies that dysfunction is scientifically nonarbitrary: "Note that [Lilienfeld and Marino's] objection concerns only the boundary between disorder and nondisorder, not the bulk of clear nonboundary cases of disorder and nondisorder" (p. 379). He continues, "What is essential is that the concept [of disorder] and its opposite can be clearly applied to a range of important cases; vagueness along the boundary is not critical and indeed is to be expected" (p. 379). And "The [HD] analysis was aimed at explaining shared judgments about a range of important cases that clearly fall on one side or the other of the boundary" (p. 379). Wakefield's language clearly suggests that in most, if not all, cases a clear natural boundary exists between dysfunction and function, although a certain degree of fuzziness often surrounds this boundary. This assertion warrants careful scrutiny.

Figure 1 displays three hypothetical distributions of phenotypes illustrating different outcomes of natural selection. For the sake of simplicity, we assume that these are population distributions measured without error. In Panel A, there is a clear threshold effect (i.e., one point of inflection); cases below a threshold value are selected for, but cases above this value are selected against. Panel B illustrates bimodality, which is typically observed in disruptive selection (Ridley, 1993). In this case, low and high values are selected for, but intermediate values are selected against. In both Panels A and B, there is a genuine natural boundary. Certain cases are unambiguously selected for and against, although there is some imprecision surrounding the boundary between selected and non-selected cases.

Contrast these two cases with that shown in Panel C. It illustrates a classic case of stabilizing selection, in which intermediate phenotypic values are selected for and increasingly extreme phenotypic values are selected against. Stabilizing selection probably accounts for the origins of many, if not most, physical and mental characteristics (Ridley, 1993), such as blood pressure, blood sugar level, and anxiety proneness. In Panel C there is no natural boundary between selected and nonselected values, only a continuous gradient of selection for intermediate values. Thus, in Panel C, there is no "there there": The distinction between dysfunction and nondysfunction is scientifically arbitrary. Consequently, in many or most cases Wakefield's (1999) assertions concerning the imprecision of the boundary are misconceived, as there often or usually is no true boundary to begin with.

Thus, Wakefield (1999) repeatedly confuses an imprecise scientific distinction surrounding a genuine boundary with an entirely arbitrary scientific distinction surrounding a nonexistent boundary. Although Wakefield intends to make the former distinction, he inevitably makes the latter distinction in the case of stabilizing selection (or any form of selection in which there is no natural demarcation point). If Wakefield were to restrict the HDA to conditions involving taxonicity, he would be forced to exclude many medical disorders (e.g., essential hypertension, diabetes mellitus) and probably many psychological disorders (e.g., most or

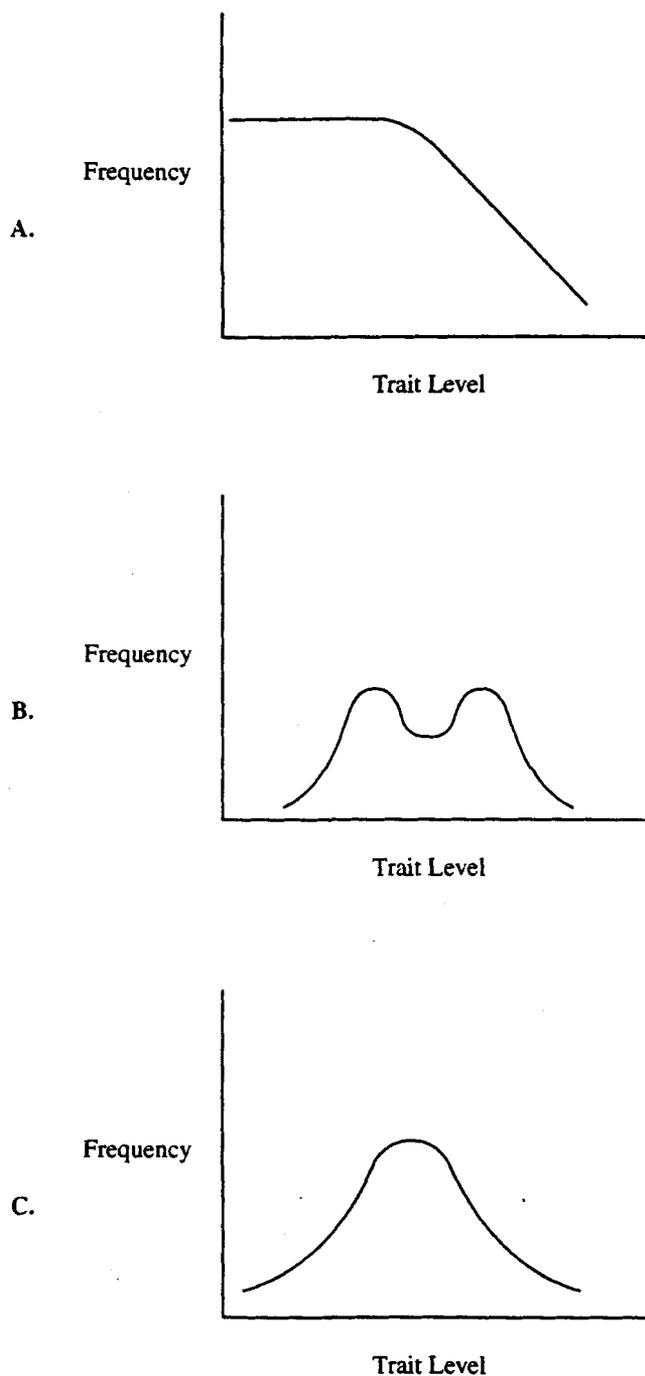


Figure 1. Three hypothetical distributions of phenotypes illustrating different outcomes of natural selection.

perhaps all anxiety, mood, and personality disorders) from the HDA, thereby falsifying the HDA by the method of conceptual analysis. In most or all of these disorders, the range of selected responses appears to be continuous, with no natural demarcation point between dysfunction and nondysfunction.

Wakefield (1999) also repeatedly assumes that extreme responses necessarily fall outside the naturally selected range. For

example, he argues on p. 389 that certain subtypes of antisocial personality disorder and social phobia may result from extremes of normal variation that produce functioning outside of the selected range. But he neglects to mention that extreme variants may themselves be maintained in the population by selection pressure. For example, a small but stable number of individuals with extreme variations of certain psychological traits may seek out and locate adaptive niches (Hutchinson, 1957) in which their trait levels are favored by natural selection. A small subset of psychopaths, for example, may be able to adopt lifestyles in which their risk-taking and fearlessness provide a net selective advantage (e.g., jet-setting entertainers who attract many mates). In addition, rare variants of certain traits may be maintained in the population by frequency-dependent selection (Ridley, 1993). Mealey (1995), for example, posited that a small number of cheaters are selected for in every human society, giving rise to a low but stable proportion of psychopaths in all cultures. Although such individuals possess extreme levels of psychopathic traits, these levels are not outside the naturally selected range.

Cultural Exaptations: The Decline and Fall of the HDA

In our previous article, we (Lilienfeld & Marino, 1995) argued that cultural exaptations (e.g., reading, music, arithmetic ability; Gould, 1991; Gould & Vrba, 1982), that is, nonselected by-products of capacities that have themselves been selected for, pose a major problem for the HDA, because they are not evolutionarily designed and consequently cannot be dysfunctions as defined Wakefield.⁴ The fact that dyslexia, amusia, and acalculia are regarded as disorders therefore appears to falsify the HDA.

Before examining Wakefield's (1999) counterarguments, we should point out that we did not contend that biological exaptations, that is, by-products of naturally selected capacities that have undergone secondary selection, are counterexamples to the HDA. Wakefield introduces unnecessary confusion on p. 380 when he quotes a passage concerning biological exaptations from our (Lilienfeld & Marino, 1995) article and then devotes an entire section of his article to a refutation of our "Biological Exaptation Objection." In this section, he criticizes our passage and explains why exaptations that have undergone secondary selection can be regarded as evolutionarily designed. Inexplicably, however, Wakefield omits a crucial sentence from our passage: "Nonetheless, a modification of Wakefield's conceptualization of dysfunction to refer to present, rather than past, design could accommodate many secondary adaptations, namely those that are currently adaptive" (Lilienfeld & Marino, 1995, p. 412). Consequently, Wakefield's subsequent nine paragraphs addressing biological exaptations are directed toward a phantom enemy and are irrelevant to our arguments.

⁴ In addition to cultural exaptations, a major challenge to the primary of natural selection as a mechanism of evolution is the Neutral Theory, which posits that most evolutionary change at the molecular level is produced by random genetic drift (Kimura, 1983). Several other non-Darwinian processes can result in evolution, including hierarchical effects at a selected level exerting a non-Darwinian effect at another level, correlated change in one feature in order to accommodate another feature (i.e., allometry), and "hitch-hiking" in which a self-promoting genetic element exerts phenotypic effects without natural selection (e.g., Gould, 1991; Gould & Lewontin, 1979).

But how does Wakefield (1999) attempt to deal with cultural exaptations, which as we noted pose a more central challenge to the HDA? He maintains that cultural exaptations are entirely consistent with the HDA, because this "analysis contains no adaptationist bias that evolutionary dysfunctions will be found at the root of all human problems" (p. 384). Specifically, Wakefield argues that our criticisms are unwarranted, because they are premised on the incorrect assumption that the harm component of the HDA must be directly produced by the dysfunction. He asserts that "There is no requirement that the harm must be identical to the failed natural function" (p. 382) and that harm can be an indirect consequence of any dysfunction.

Wakefield (1999) suggests that our "confusion [concerning this issue] is perhaps understandable" (p. 383). We agree: Wakefield (1992a) asserted that "Because natural selection is the only known means by which an effect can explain a naturally occurring mechanism that provides it, evolutionary explanations presumably underlie all correct ascriptions of natural functions" (p. 383). It seems clear that Wakefield did not regard nonselected exaptations as relevant to the HDA, and his current disclaimer concerning hyperadaptationist bias is puzzling.

These apparent contradictions notwithstanding, what happens to the validity of the HDA when the harm and dysfunction components are dissociated? One of the strengths of the HDA (as we understood it) was that the harms experienced by the organism (e.g., extreme anxiety) were tied fairly directly to dysfunctions of naturally selected mechanisms (e.g., the fight-flight system). But now that harm and dysfunction are dissociated, virtually any harmful consequence can be a disorder, because virtually any harm can be an indirect result of some dysfunction somewhere in the body.

To illustrate this point, we examine Wakefield's (1999) argument concerning the status of acalculia as a disorder:

The attribution of disorder of acalculia is based on a line of reasoning roughly as follows: (a) inability to learn to calculate is a significant harm; (b) the brain was not designed specifically to enable people to learn to calculate; (c) however, when all of a person's brain systems are functioning as they were designed to function, a side-effect is that the person can learn to calculate; (d) therefore, inability to learn to calculate (despite conducive environmental and motivational circumstances) is caused by some . . . underlying brain system failing to function as it was designed to function, and is a disorder. (p. 383)

To evaluate the validity of this line of reasoning, we suggest that the reader perform the following three steps: (a) substitute any or all of the following words/phrases for "learn to calculate" in Steps (a) through (d): "cook," "bathe," and "dress oneself"; (b) substitute "a person's arms" for "all of a person's brain systems" in Step (c) and "some underlying brain system" in Step (d); and (c) substitute "they" for "it." The HDA has now degenerated into absurdity: Individuals who lose the use of their arms through a degenerative disease or injury should be considered to suffer from "cooking disorder," "bathing disorder", and "dressing disorder" according to the HDA, as all of these conditions cause harm to the vast majority of affected individuals. Similarly, individuals who lose the use of their eyes, legs, or arms as a result of disease or injury should be considered to suffer from "driving disorder" according to the HDA. The number of falsifying counterexamples to the HDA is enormous, because as Wakefield (1992a) himself observed, "the

list of possible harms is potentially endless" (p. 381). The fact that inability to cook, bathe, dress oneself, and drive a car are not considered disorders suggests that something is seriously wrong with the HDA. It should be noted that because the HDA posits that harm and dysfunction are singly necessary and jointly sufficient for disorder, the HDA cannot be rescued by arguing that only certain harmful dysfunctions are disorders.

Indeed, Wakefield's (1999) Conceptual Experiment 2 affords an opportunity to directly compare the validity of the HDA and RA. According to Wakefield, the HDA predicts that failure to learn to read due to distractions, lack of motivation, and so on, should not be considered a disorder, whereas failure to learn to read due to a dysfunction should. Moreover, because the HDA asserts that harm is not necessarily a direct consequence of dysfunction, it predicts that failure to learn to read due to disease-related blindness (a dysfunction according to Wakefield) should be considered a disorder, which it is not. Nor is the failure to learn to speak as a result of congenital deafness regarded as a disorder. Once again, the HDA is clearly falsified.

Why do individuals not perceive inability to cook, bathe oneself, and so on, as disorders, particularly when they are not a result of brain damage? The reason, we suggest, is that individuals perceive these capacities as relatively remote consequences of other bodily and mental functions. As a result, when these capacities malfunction, individuals do not perceive such malfunction per se as reflecting something wrong with the body or mind but rather as an indirect consequence of another bodily or mental problem. In contrast, dyslexia, acalculia, and amusia are typically perceived as disorders because they are viewed as bearing a relatively direct link to brain functioning. The RA correctly predicts these consensual judgments; the HDA does not.

Mismatch Theory and Disorder

A number of authors have recently argued that many physical and mental disorders result from a mismatch between evolutionarily designed functioning and novel environments (Glantz & Pierce, 1989; Nesse & Williams, 1994). Because the pace of cultural evolution has outstripped the pace of biological evolution in many domains, certain physical and mental reactions that were adaptive in the Pleistocene epoch may give rise to disorders in modern (i.e., technological) environments. If so, the disorders resulting from such mismatches would pose a problem for the HDA, because such disorders are produced by systems that are functioning as designed.

Myopia, for example, appears in many cases to be produced by reading or exposure to other close-up stimuli early in life. Interestingly, myopia appears to be rare in technologically undeveloped societies but common in modern environments (Curtin, 1988; Young et al., 1969). In early development, the brain apparently stimulates retinal growth whenever it receives a signal from the retina that the visual image is blurred. The contemporary practice of exposing young children to written materials produces a blurred visual image and may result in the excessive retinal growth characteristic of myopia (Bock & Widdows, 1990; Nesse & Williams, 1994). In such cases, the eye is functioning as designed given the retinal input it has received. Weder and Schork (1994) posited that essential hypertension typically results from mismatches between the designed propensity of the cardiovascular system to increase

blood pressure in response to growth spurts in childhood and adolescence and the increased body size of modern children and adolescents produced by improvements in nutrition. Essential hypertension, like myopia, is rare in technologically undeveloped societies (Weder & Schork, 1994).

Several authors have argued that pathological obesity is also a consequence of current dietary excesses. Our hunger systems were probably designed to be efficient at storing calories during periods of short-term surplus, because periods of prolonged famine were frequent (Konner, 1982). The environments of modern Western society, which are often characterized by abundant supplies of food at all times, therefore contribute to obesity in many individuals. A similar argument has been made for many cases of adult-onset diabetes. Interestingly, efforts to combat high levels of malnutrition among the Pima Indians of Arizona, who had become accustomed over many generations to low levels of food intake, resulted in marked increases in the prevalence of diabetes and obesity (Neel, 1962).

A number of psychological disorders may also result from mismatches between evolutionary design and novel environments. Some specific phobias can be viewed as reactions that were adaptive prior to certain technological developments. Blood phobia, for example, involves a set of pronounced parasympathetic reactions (e.g., heart rate and blood pressure decreases) designed to minimize blood loss (Barlow, 1988). Such extreme reactions were probably adaptive prior to the development of bandages and coagulants, but are unnecessary in most modern environments (Lilienfeld & Marino, 1995).

Moffitt (1993) hypothesized that many cases of adolescent conduct disorder are a consequence of a "maturity gap" produced by modern environmental changes in interaction with evolved propensities toward mimicry. Specifically, recent improvements in nutrition have decreased the average age of puberty, while recent advances in technology and education have increased the average age at which adolescents enter the work force. Consequently, Moffitt argued, today's adolescents are "chronological hostages of a time warp" (p. 687) in which they are biologically mature but incapable of assuming adult status. During this time warp, they mimic the actions of peers in an effort to simulate the behaviors of adults and are therefore susceptible to deviant peer influences and to the development of conduct disorder.

Wakefield (1999) acknowledges that mismatches between evolutionary design and modern environments can produce maladaptive behavior but maintains that such behavior is not considered disordered because "that is how humans are designed" (p. 385). But as noted above, a number of such mismatches probably do result in consensual disorders. Moreover, Wakefield's argument that "design failure rather than current maladaptiveness is the final arbiter in judgments of disorder" (p. 385) is unconvincing. To see why, imagine a modification of Conceptual Experiment 7 in which modern nutritional resources had deteriorated to the point that the diet of individuals with a variant folic metabolism gene could not be modified, so that virtually all such individuals developed severe symptoms of pernicious anemia and required treatment. Such individuals would almost certainly be regarded as disordered, despite the fact that their symptoms are a consequence of recent nutritional changes in interaction with a naturally selected genetic variant. This judgment is inconsistent with the HDA but consistent with the RA, which posits that individuals are judged as disordered

if they are perceived as experiencing a problem with the body that requires amelioration.

Defects Versus Defenses

The HDA neglects the distinction between *defects* and *defenses* (Nesse & Williams, 1994) and essentially regards all disorders as defects. Defects are malfunctions of physical or mental systems that serve no useful role. The plaques and tangles of Alzheimer's dementia, for example, are nonadaptive by-products of a disease process. Defenses, in contrast, are naturally selected responses (e.g., coughing, vomiting) designed to protect the organism against external threats. If some consensual disorders consist of defenses, this would falsify the HD analysis, because such disorders would not involve failures of systems to function as designed.

Wakefield (1999) takes issue with our claim (Lilienfeld & Marino, 1995) that the primary features of a flu are adaptive defenses against infection and are therefore not dysfunctions. He argues that (a) sneezing, coughing, and fever are symptoms of flu, rather than disorders per se (which, contrary to Wakefield's comments, we never disputed; see Lilienfeld & Marino, 1995, p. 415) and (b) flu involves a dysfunction and thus does not represent a counterexample to the HDA. But where is the dysfunction in the flu? Wakefield asserts that "When the accumulation of mucus in the respiratory tract due to a flu causes sneezing and coughing, it is not the sneezing and coughing in and of themselves but the underlying flu or respiratory blockage that is the disorder" (p. 391).

This assertion bears careful examination. First, the underlying flu cannot be the disorder according to the HDA, because the flu itself is not a failure of a system to perform its designed function and therefore is not a dysfunction. Second, the claim that the respiratory blockage is the disorder is unwarranted. The shedding of mucus that produces this blockage is itself an adaptive defense against viral infection. Moreover, the respiratory blockage produced by mucal secretion illustrates precisely the kind of functional hierarchy that Wakefield (1999) himself discusses on p. 394 (see also Klein, 1978): When one or more systems (e.g., the mucous membranes) react in a designed fashion to threat (e.g., a virus) and temporarily override the functioning of another system (e.g., the respiratory tract), "there is no dysfunction because the overall system is performing as designed" (Wakefield, 1999, p. 394). Ironically, Wakefield's own reasoning suggests that a flu is a coordinated set of defenses against infection, rather than a dysfunction.

Wakefield's (1999) claim on p. 393 that the edema resulting from a sprained ankle is not a disorder is equally unconvincing. Wakefield does not dispute the contention that edema is a consensual disorder. Instead, he reviews evidence indicating that edema is a designed response to tissue damage and concludes that edema is neither a dysfunction nor a disorder. Wakefield's reasoning is entirely tautological: He (a) argues that consensual disorders are design failures, (b) contends that at least one consensual disorder, namely edema, is not a design failure, and (c) concludes that edema must not be a disorder after all! But the fact that edema is a designed response actually refutes the HDA according to the method of conceptual analysis. If Wakefield were to respond that consensual judgments of disorder are sometimes incorrect, then the HDA would become essentially unfalsifiable. When presented

with any counterexamples to the claim that consensual disorders involve dysfunctions, Wakefield could, as in the case of edema, simply maintain that these specific consensual judgments are mistaken.

Flu and edema aside, there is ample evidence that a number of consensual disorders involve defenses against threat.

First, many skin disorders are defensive reactions against danger. Calluses and corns, including those that are extremely painful or disfiguring, are adaptive responses of the epidermis to repeated friction. They buffer the skin against mechanical injury and decrease the likelihood of pathogen invasion (Nesse & Williams, 1994). Skin disorders characterized by desquamation (shedding), such as exfoliative dermatitis, provide another example. Because "the stratum corneum serves as a reservoir for exogenous toxic and nontoxic agents . . . the desquamation process may be a means by which the skin and body can rid itself of toxic agents that build up within the stratum corneum reservoir" (Fitzpatrick, Eisen, Wolff, & Freedberg, 1993, p. 245).

Second, as Nesse and Williams (1994) noted, "we know that the system that gives rise to allergy is a defense" (p. 159) and that "allergy is not an extreme action of some normally behaved system with an obvious function" (p. 160). The immunoglobulin-E (IgE) system appears to exist solely to trigger allergic reactions. When certain foreign substances invade the body, IgE is produced. IgE binds to mast cells, which become primed for the allergen's return. When the allergen reappears, mast cells exude substances that attack adjacent cells, draw white blood cells to the allergen, and activate smooth muscle, thereby producing asthma. Allergic reactions expel toxic substances from the body efficiently by means of eye watering, mucal secretion, sneezing, coughing, vomiting, and diarrhea, all of which are defenses against the allergen (Nesse & Williams, 1994).

Third, the symptoms of heat exhaustion appear to be a consequence of the body's adaptive efforts at temperature regulation. These symptoms (e.g., nausea, low blood pressure, cramps, vertigo) "are caused by sodium chloride depletion that accompanies profuse sweating. These symptoms can occur because temperature regulation has priority over the maintenance of salt and water balance despite their severe depletion" (Fitzpatrick et al., 1993, p. 411). The overriding of salt and water balance again illustrates the operation of functional hierarchies as described by Wakefield (1999) and strongly implies that the symptoms of heat exhaustion are not dysfunctions.

Fourth, staphylococcal poisoning (and several other forms of food poisoning) is characterized by vomiting, diarrhea, and fever, which are adaptive attempts to eliminate bacterial infection (Berkow & Fletcher, 1992). The other common symptoms of staphylococcal poisoning (e.g., headaches, abdominal cramps) are unpleasant consequences of infection but are not dysfunctions.

Thus, Wakefield's (1999) assertion that "designed reactions to threat . . . are not in themselves considered disorders" (p. 397) is refuted, because a number of physical disorders consist of adaptive defenses rather than dysfunctions. Indeed, Wakefield overlooks the critical point that these defenses are designed to *prevent* dysfunctions, especially those that could endanger the long-term health of the organism. By implicitly treating disorders characterized by defenses as defects, the HDA, although intended to provide an evolutionary account of disorder, is actually antievolutionary in both its substance and spirit. It should also be noted that the fact

that flu, allergy, and so on are considered disorders is explained by the RA, but not by the HDA. In all of these conditions, individuals perceive that there is something wrong with the body that needs to be fixed, even though no dysfunction is present.

The Assimilative Nature of Personality Traits: Psychological Defenses and Emotional Disorders

Wakefield (1999) disputes our contention that many emotional disorders (e.g., panic disorder) are naturally selected defenses against threat. He acknowledges that the defensive reactions comprising these conditions (e.g., panic attacks) are topographically normal but argues that these reactions occur in response to stimuli (e.g., relatively innocuous threats) that are outside the range for which these reactions were designed. Therefore, he concludes, these responses are dysfunctional.

But the HDA neglects the assimilative nature of personality traits (Tellegen, 1991; see also Bowers, 1973 and Wachtel, 1977). Allport's (1961) classic statement that personality traits have "the capacity to render many stimuli functionally equivalent" (p. 347) and Murray's (1938) distinction between alpha and beta press underscore the fact that traits influence how individuals construe and interpret stimuli (or, in Piagetian terms, lead individuals to "assimilate" ambiguous information into schemas). The assimilative nature of traits bears important implications for the HDA. Many emotional disorders are probably consequences of psychological systems reacting in a designed fashion to subjectively threatening input, rather than dysfunctions.

Take, for example, the construct of anxiety sensitivity (AS), which is an antecedent and perhaps a risk factor for panic disorder (Lilienfeld, Jacob, & Turner, 1993; McNally, 1994). Individuals with high levels of AS believe that anxiety symptoms (e.g., rapid heartbeat) will lead to catastrophic consequences (e.g., heart attack). As a consequence, they subjectively perceive these symptoms to be much more threatening than do other individuals. From this perspective, the panic (i.e., fight-flight) system of the panic disorder patient with elevated AS is reacting precisely as it was designed to act: It is receiving a signal that danger is imminent and becomes activated in response to this signal. The fight-flight reactions of panic patients are inappropriate in the sense that they occur more frequently and in response to a broader range of stimuli than those of other individuals, but they are not dysfunctional in Wakefield's (1999) sense.

Similar arguments can be made regarding the etiology of many mood disorders. High levels of negative affectivity (NA; Watson & Clark, 1984), for example, have been found to predate the onset of major depressive episodes (Hirschfeld et al., 1989). Individuals with elevated NA tend to construe minor setbacks and losses as disastrous occurrences (Watson & Clark, 1984) and to perceive ambiguous stimuli as more threatening compared with other individuals (Eysenck & Mathews, 1987). Therefore, it can be argued that the mood system of the high NA individual who becomes clinically depressed following the loss of a relationship or job is not itself dysfunctional. Instead, this system is reacting as designed given the threatening input it has received.

It should be noted that because many or most personality traits are probably dimensional (i.e., nontaxonic) in nature (Eysenck, 1994; Widiger, 1993), Wakefield could not plausibly respond to these examples by maintaining that high levels of personality traits

typically constitute dysfunctions. As noted earlier, the HDA requires that every case of dysfunction (and thus every case of disorder) is characterized by a genuine point of demarcation separating designed from nondesigned functioning.

Wakefield (1999, p. 395) attempts to refute our arguments that panic disorder and major depression often represent normal defenses expressed in inappropriate situations by presenting several examples of medical disorders (e.g., cancer, cardiac arrhythmia) characterized by normal physical responses manifested under inappropriate conditions. But few, if any, of these conditions involve a system reacting in a designed fashion to extreme input. To the contrary, several of these conditions (e.g., autoimmune disorders; Nesse & Williams, 1994) appear to involve a system overreacting to essentially normal input. Thus, the analogy to the emotional disorders discussed here is misleading.⁵

Conclusion: Essentialism and Biological Concepts

According to the HDA, dysfunction of a naturally selected system lies at the core of all disorders, both psychological and medical. We have demonstrated, however, that (a) some dysfunctions (i.e., physiological breakdowns leading to failures of many cultural exaptations) are not regarded as disorders and (b) some conditions regarded as disorders (i.e., mismatches between naturally selected responses and novel environments, defenses against threat) are not dysfunctions. Yet both sets of counterexamples are explained by the RA: Failures of many cultural exaptations (e.g., cooking) do not lead individuals to conclude that there is something wrong with the body or mind that needs to be fixed, whereas design–environment mismatches and defenses against extreme threat do. We conclude that Wakefield's (1999) criticisms of the RA do not withstand careful scrutiny and that the HDA has been convincingly falsified. We readily acknowledge, however, that certain details of our RA will require elaboration and that the question of what factors lead individuals to make the inference that there is something wrong with the body or mind that needs to be fixed remains an important area for future research.

Wakefield (1999) argues on p. 397 that humans' propensity toward psychological essentialism reflects the fact that natural concepts are often underpinned by deeper causes. We agree. But it would be remarkable, in our view, if widely shared judgments of disorder mapped perfectly or almost perfectly onto one, and only one, latent cause in nature—namely, the failure of a naturally selected system to perform its designed function. Although the heuristics identified by social cognition researchers (e.g., representativeness; Tversky & Kahneman, 1974) tend to be helpful in organizing and categorizing the complex information of the real world, they often lead us to go beyond the data and to perceive underlying commonalities among fundamentally unrelated stimuli (Nisbett & Ross, 1980). Wakefield is correct that many disorders involve evolutionary dysfunctions, but he is incorrect that such dysfunctions can provide a definitive criterion for distinguishing disorder from nondisorder.

How well has essentialism fared in clarifying the meaning of other natural concepts? To answer this question, it may be helpful to examine the status of essentialistic analyses in resolving controversies regarding two other concepts relevant to evolutionary theory: species and life. The extensive hybridization and morphological overlap among putatively different "species" (Mayr, 1982)

have led most biologists to conclude that attempts to provide necessary and sufficient criteria for the concept of species are futile. Levin (1979), for example, suggested that "the search for hidden likenesses is unlikely to yield a unifying species concept" and that the species concept is "a mental abstraction which orders clusters of diversity in multidimensional space" (p. 381). Hull (1976) noted that "initially it was thought that the names of all species could be defined by sets of essentialistic traits" but that "no matter how hard they tried, taxonomists could rarely find sets of traits which divided living organisms into neat little packets . . . Our inability to distinguish most species by sets of necessary and sufficient conditions follows from evolutionary theory . . ." (p. 180; see Eberhardt & Randall, 1997, for a similar critique of essentialistic concepts of race).

With regard to the concept of life, Medawar and Medawar (1983) noted that

A great many nonbiologists believe that animated and contentious discussions of the definition of "life" are a principal preoccupation of institutes and university departments of biology. In reality, the subject is not mentioned at all, except to disparage the rather simple-minded people who believe that an agreed-upon definition of life will lead to a better comprehension of biology. . . . A hunger for definitions is very often a manifestation of a deep-seated belief . . . that all words have an inner meaning that patient reflection and research will make clear . . . indeed, amateurs will sometimes [ask]: "What is the true meaning of the word 'life'?" There is no true meaning. There is a usage that serves the purposes of working biologists well enough. (pp. 66–67)

We certainly do not wish to imply that Wakefield (1999) and others who have proposed essentialistic analyses of disorder are "simple-minded" or to summarily dismiss the possibility that the disorder concept may ultimately yield to an essentialistic account that remedies the shortcomings of the HDA. But the comments of Levin (1979), Hull (1976), and the Medawars (1983) should give us pause for two reasons.

First, the fact that two other biological concepts relevant to evolutionary theory have proven intractable to essentialistic analyses should perhaps lead those in the field to question the assumption that the concept of disorder will necessarily be amenable to such an analysis. Second, the fact that biologists have made substantial progress in many domains without developing an explicit operationalization of either species or life should perhaps also lead us to question the assumption that the absence of such an operationalization will necessarily impede progress in psychopathology research. Wakefield (1996) argued that the "valid discrimination of disorder from nondisorder" should be the "essential intellectual goal" (p. 647) of the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; American Psychiatric Association, 1994) and other diagnostic manuals. But if biologists do not lose

⁵ Note that our analysis of emotional disorders as defenses does not conform to Wakefield's (1999) example of "dysfunctional interactions" (p. 386) between two mechanisms. In Wakefield's neurotransmitter example, the relations between the two mechanisms are interactive (i.e., synergistic). But in the examples of emotional disorders presented here, the responses judged as disordered are a consequence of main effects, not interactions: One system (e.g., the panic system) receives extreme input from another system and responds precisely as it was designed to respond.

sleep over the boundaries of their subject matter, why must psychologists and psychiatrists be any different?

References

- Allport, G. (1961). *Pattern and growth in personality*. New York: Holt, Reinhart, & Winston.
- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- Barlow, D. H. (1988). *Anxiety and its disorders: The nature and treatment of anxiety and panic*. New York: Guilford Press.
- Berkow, R., & Fletcher, A. J. (1992). *The Merck manual of diagnosis and therapy*. Rahway, NJ: Merck.
- Bock, G. R., & Widdows, K. (1990). *Myopia and the control of eye growth*. Chichester, England: Wiley.
- Bowers, K. S. (1973). Situationism in psychology: An analysis and critique. *Psychological Review*, 80, 307–336.
- Campbell, E. J. M., Scadding, J. G., & Roberts, R. S. (1979). The concept of disease. *British Medical Journal*, 2, 757–762.
- Cantor, N., Smith, E. E., French, R. D., & Mezzich, J. (1980). Psychiatric diagnosis as prototype categorization. *Journal of Abnormal Psychology*, 89, 181–193.
- Cohen, H. (1981). The evolution of the concept of disease. In A. L. Caplan, H. T. Engelhardt, Jr., & J. J. McCartney (Eds.), *Concepts of health and disease: Interdisciplinary perspectives* (pp. 209–220). Reading, MA: Addison-Wesley.
- Curtin, B. J. (1988). *The myopias*. Philadelphia: Harper & Row.
- Dawes, R. M. (1994). *House of cards: Psychology and psychotherapy built on myth*. New York: Free Press.
- Eberhardt, J. L., & Randall, J. L. (1997). The essential notion of race. *Psychological Science*, 8, 198–203.
- Eysenck, H. J. (1994). Normality–abnormality and the three-factor model of personality. In S. Strack & M. Lorr (Eds.), *Differentiating normal and abnormal personality* (pp. 3–25). New York: Springer.
- Eysenck, M. W., & Mathews, A. (1987). Trait anxiety and cognition. In H. J. Eysenck & I. Martin (Eds.), *Theoretical foundations of behavior therapy* (pp. 197–216). New York: Plenum Press.
- Fitzpatrick, T. B., Eisen, A. Z., Wolff, K., Freedberg, I. M., & Austen, K. F. (1993). *Dermatology in general medicine: Vol. 1* (4th ed.). New York: McGraw-Hill.
- Glantz, K., & Pearce, J. (1989). *Exiles from Eden: Psychotherapy from an evolutionary perspective*. New York: Norton.
- Gorenstein, E. E. (1984). Debating mental illness: Implications for science, medicine, and social policy. *American Psychologist*, 39, 50–56.
- Gould, S. J. (1991). Exaptation: A crucial tool for evolutionary analysis. *Journal of Social Issues*, 47, 43–65.
- Gould, S. J., & Lewontin, R. C. (1979). The Spandrels of San Marcos and the Panglossian paradigm: A critique of the adaptationist programme. *Proceedings of the Royal Society of London*, 205, 581–598.
- Gould, S. J., & Vrba, E. S. (1982). Exaptation: A missing term in the science of form. *Paleobiology*, 8, 4–15.
- Hirschfeld, R. M. A., Klerman, G. L., Lavori, P., Keller, M., Griffith, P., & Coryell, W. (1989). Premorbid personality assessments of first onset of major depression. *Archives of General Psychiatry*, 46, 345–350.
- Hull, D. L. (1976). Are species really individuals? *Systematic Zoology*, 25, 174–191.
- Hutchinson, G. E. (1957). Concluding remarks. *Cold Spring Harbor Symposium on Quantitative Biology*, 22, 415–427.
- James, W. (1890). *Principles of psychology* (Vol. 1). New York: Holt.
- Kazdin, A. E. (1983). Psychiatric diagnosis, dimensions of dysfunction, and child behavior therapy. *Behavior Therapy*, 14, 73–99.
- Kendell, R. E. (1975). The concept of disease and its implications for psychiatry. *British Journal of Psychiatry*, 127, 305–315.
- Kendell, R. E. (1986). What are mental disorders? In A. M. Freedman, R. Brotman, I. Silverman, & D. Hutson (Eds.), *Issues in psychiatric classification: Science, practice and social policy* (pp. 23–45). New York: Human Sciences Press.
- Kimura, M. (1983). *The neutral theory of molecular evolution*. Cambridge, England: Cambridge University Press.
- Kirmayer, L. J. (1994). Rejoinder to Professor Wakefield. In S. A. Kirk & S. D. Einbinder (Eds.), *Controversial issues in mental health* (pp. 17–20). Boston: Allyn & Bacon.
- Klein, D. F. (1978). A proposed definition of mental illness. In R. L. Spitzer & D. F. Klein (Eds.), *Critical issues in psychiatric diagnosis* (pp. 41–71). New York: Raven Press.
- Komatsu, L. K. (1992). Recent views of conceptual structure. *Psychological Bulletin*, 112, 500–526.
- Konner, M. (1982). *The tangled wing: Biological constraints on the human spirit*. New York: Holt.
- Kräupl Taylor, F. (1971). A logical analysis of the medicopsychological concept of disease. *Psychological Medicine*, 1, 356–364.
- Levin, D. A. (1979). The nature of plant species. *Science*, 204, 381–384.
- Liljenfeld, S. O. (1995). Reply to Maltzman's "Why alcoholism is a disease." *Journal of Psychoactive Drugs*, 27, 287–291.
- Liljenfeld, S. O., Jacob, R. G., & Turner, S. M. (1993). Anxiety sensitivity: An examination of theoretical and methodological issues. *Advances in Behaviour Therapy and Research*, 15, 147–183.
- Liljenfeld, S. O., & Marino, L. (1995). Mental disorder as a Roschian concept: A critique of Wakefield's "harmful dysfunction" analysis. *Journal of Abnormal Psychology*, 104, 411–420.
- Liljenfeld, S. O., Waldman, I. D., & Israel, A. C. (1994). A critical examination of the use of the term "comorbidity" in psychopathology research. *Clinical Psychology: Science and Practice*, 1, 71–83.
- Lykken, D. T. (1995). *The antisocial personalities*. Hillsdale, NJ: Erlbaum.
- Mayr, E. (1982). *The growth of biological thought: Diversity, evolution, and inheritance*. Cambridge, MA: Belknap Press.
- McNally, R. J. (1994). *Panic disorder: A critical analysis*. New York: Guilford Press.
- Mealey, L. (1995). The sociobiology of sociopathy: An integrated evolutionary model. *Behavioral and Brain Sciences*, 18, 523–541.
- Medawar, P. B., & Medawar, J. S. (1983). *Aristotle to zoos: A philosophical dictionary of biology*. Cambridge: Harvard University Press.
- Medin, D. (1989). Concepts and conceptual structure. *American Psychologist*, 44, 1469–1481.
- Medin, D., & Ortony, A. (1989). Psychological essentialism. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 179–195). New York: Cambridge University Press.
- Meehl, P. E., & Golden, R. (1982). Taxometric methods. In P. C. Kendall & J. N. Butcher (Eds.), *Handbook of research methods in clinical psychology* (pp. 127–181). New York: Wiley.
- Mervis, C. (1980). Category structure and the development of categorization. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension: Perspectives from cognitive psychology, linguistics, artificial intelligence, and education* (pp. 279–307). Hillsdale, NJ: Erlbaum.
- Moffitt, T. E. (1993). Adolescent-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review*, 100, 674–701.
- Murray, H. A. (1938). *Explorations in personality*. New York: Oxford.
- Neel, J. V. (1962). Diabetes mellitus: A "thrifty" genotype rendered detrimental by "progress"? *American Journal of Human Genetics*, 14, 353–362.
- Neisser, U. (1979). The concept of intelligence. *Intelligence*, 3, 217–227.
- Nesse, R. M., & Williams, G. C. (1994). *Why we get sick: The new science of Darwinian medicine*. New York: Times Books.
- Nisbett, R., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice Hall.
- Ridley, M. (1993). *Evolution*. Boston: Blackwell Scientific Publications.

- Rosch, E. R. (1973). Natural categories. *Cognitive Psychology*, 4, 328–350.
- Rosch, E. R., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 7, 573–605.
- Rothbart, M., & Taylor, M. (1992). Social categories and social reality. In G. R. Semin & K. Fiedler (Eds.), *Language, interaction, and social cognition* (pp. 11–36). Newbury Park, CA: Sage.
- Sedgwick, P. (1982). *Psycho politics*. New York: Harper & Row.
- Tellegen, A. (1991). Personality traits: Issues of definition, evidence, and assessment. In D. Cicchetti & W. Grove (Eds.), *Thinking clearly about psychology: Essays in honor of Paul Everett Meehl* (pp. 10–35). Minneapolis, MN: University of Minnesota Press.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124–1131.
- Wachtel, P. (1977). *Psychoanalysis and behavior therapy: Toward an integration*. New York: Basic Books.
- Wakefield, J. C. (1992a). The concept of mental disorder: On the boundary between biological facts and social values. *American Psychologist*, 47, 373–388.
- Wakefield, J. C. (1992b). Disorder as harmful dysfunction: A conceptual critique of *DSM-III-R*'s definition of mental disorder. *Psychological Review*, 99, 232–247.
- Wakefield, J. C. (1993). Limits of operationalization: A critique of Spitzer and Endicott's (1978) proposed operational criteria for mental disorder. *Journal of Abnormal Psychology*, 102, 160–172.
- Wakefield, J. C. (1996). *DSM-IV*: Are we making diagnostic progress? *Contemporary Psychology*, 41, 646–652.
- Wakefield, J. C. (1999). Evolutionary versus prototype analyses of the concept of disorder. *Journal of Abnormal Psychology*, 108, 374–399.
- Watson, D., & Clark, L. A. (1984). Negative affectivity: The disposition to experience aversive emotional states. *Psychological Bulletin*, 96, 465–490.
- Weder, A. B., & Schork, N. J. (1994). Adaptation, allometry, and hypertension. *Hypertension*, 24, 145–156.
- Widiger, T. A. (1993). The DSM-III-R categorical personality disorder diagnoses: A critique and an alternative. *Psychological Inquiry*, 4, 75–90.
- Young, F. A., Leary, G. A., Baldwin, W. R., West, D. C., Box, R. A., & Harris, E. (1969). The transmission of refractive errors within Eskimo families. *American Journal of Optometry and Physiological Optics*, 46, 676–685.

Received February 11, 1997

Revision received August 1, 1997

Accepted August 4, 1997 ■

Low Publication Prices for APA Members and Affiliates

Keeping you up-to-date. All APA Fellows, Members, Associates, and Student Affiliates receive—as part of their annual dues—subscriptions to the *American Psychologist* and *APA Monitor*. High School Teacher and International Affiliates receive subscriptions to the *APA Monitor*, and they may subscribe to the *American Psychologist* at a significantly reduced rate. In addition, all Members and Student Affiliates are eligible for savings of up to 60% (plus a journal credit) on all other APA journals, as well as significant discounts on subscriptions from cooperating societies and publishers (e.g., the American Association for Counseling and Development, Academic Press, and Human Sciences Press).

Essential resources. APA members and affiliates receive special rates for purchases of APA books, including the *Publication Manual of the American Psychological Association*, and on dozens of new topical books each year.

Other benefits of membership. Membership in APA also provides eligibility for competitive insurance plans, continuing education programs, reduced APA convention fees, and specialty divisions.

More information. Write to American Psychological Association, Membership Services, 750 First Street, NE, Washington, DC 20002-4242.