

Running head: HARMFUL DYSFUNCTION, MENTAL DISORDERS AND DISEASE

The Harmful Dysfunction Analysis and the Differentiation

Between Mental Disorder and Disease

Gregg R. Henriques, Ph.D.

Department of Psychiatry

University of Pennsylvania

Henriques, G. R. (2002). The harmful dysfunction analysis and the differentiation between mental disorder and disease. *Scientific Review of Mental Health Practice, 1* (2), 157-173.

Correspondence regarding this article should be sent to: Gregg Henriques, Ph.D., Science Center Room 2029, 3535 Market Street, University of Pennsylvania, Philadelphia, PA 19104-2648. Email address is henri@landru.cpr.upenn.edu.

Abstract

Wakefield's Harmful Dysfunction Analysis (HDA) for distinguishing disorders from nondisorders has received much attention in the literature. Although the analysis has many strengths, Wakefield (1999a; 1999b) fails to appropriately capture the nature of the disorder construct thereby leading to much confusion. A solution is offered suggesting disorder can be thought of as a utilitarian construct. When viewed in this light, the HDA offers an excellent and useful definition of disease for medicine. However, the HDA fails as a useful definition for mental disorders because it contains a greedily reductionistic error that suggests all mental disorders are reducible to biological theory. An alternative way of conceptualizing mental disorders is offered and it is suggested that the HDA's success in defining disease provides an important piece that allows mental health scientists begin to answer which mental disorders are akin to medical diseases and which mental disorders are not.

keywords: psychiatry, psychopathology, evolution, illness, natural selection,
essentialist concept

The Harmful Dysfunction Analysis and the Differentiation
Between Mental Disorder and Disease

The debate over what constitutes a legitimate definition of mental disorder has raged since the earliest days of psychology and psychiatry. Most in psychology and psychiatry are aware of the challenges to the entire notion of mental illness offered by Szasz (1974) and the heated debates regarding the classification of homosexuality several decades ago. Conflicting opinions about whether or not phenomena such as premenstrual dysphoria (Ginsburg & Carter, 1987), anti-social behavioral tendencies (Koski & Mangold, 1993), and binge eating (Hetherington, 1993) should be classified as disordered continue today. The debates about the nature of mental disorders have been fierce because the concept is such an important one. Whether or not a condition or individual is disordered carries a host of socio-political implications, such as health care treatment and assistance, social control, stigmatization, media attention, and the dispersal of resources for research.

Over the past decade, Jerome C. Wakefield has emerged as a key voice in this arena with his proposal for explaining the disorder concept, called the Harmful Dysfunction Analysis ([HDA]; 1992a, 1992b, 1997a, 1999a, 1999b). The HDA proposes that the concept of disorder consists of two equally important components, a socially determined harm and a scientifically determined dysfunction of an internal mechanism, whereby dysfunction is defined in terms of the mechanism failing to perform its naturally selected function. Because of its inclusion of both components, the HDA can be considered a hybrid that combines previously offered biologically scientific approaches (e.g., Kendall, 1975; Scadding, 1990) with previously offered social value approaches (e.g., Sedgwick, 1982). In short, Wakefield's analysis has generated an AND out of an historical OR.

The HDA has drawn a substantial amount of attention and recently (1999) a special issue of the Journal of Abnormal Psychology was devoted to examining the concept. Both supporters and critics in the special issue generally agreed that Wakefield's combination of scientific judgment with social value was a strong move forward. A second strength of the HDA is that it clearly anchors the concept of disorder to scientific investigations. Third, Wakefield (1992a, 1999a) effectively articulates that evolutionary analyses are crucial to analyses of biological function and dysfunction and that, in many instances, we can distinguish ordered from disordered biological functioning. Finally, Wakefield demonstrates that many, if not all, disorder attributions in medical domains other than psychiatry strongly coincide with notions of dysfunction derived either explicitly or implicitly from evolutionary theory. For example, a heart attack is an example of a dysfunction because the heart was fashioned via evolutionary processes to circulate blood throughout the body. Broken bones, cancers, and strokes are also clear examples of harmful dysfunctions readily identified by the HDA. Being shorter than average, fevers in response to infection, wrinkles due to aging, and pain from a broken bone are not dysfunctions, as these occurrences do not involve breakdowns of evolved mechanisms. Taken together, these positives point to Wakefield's framework being a significant and important advance over previous definitions of disorder or disease.

Despite these advantages, there are two major problems with the HDA in its current form. First, although Wakefield should be commended for recognizing that socially constructed values and scientific judgments both make up the concept of disorder, I argue that his hybridization of the two components has ultimately been unsuccessful and that Wakefield fails to appropriately capture the nature of the disorder construct. Because of this failure, much confusion results regarding what the HDA is purporting to explain. In contrast to Wakefield, I

argue that neither disorder nor dysfunction are “essentialistic concepts,” that the HDA is not falsifiable in its current form, and that Wakefield confusingly vacillates between arguing that the HDA is a prescription for how disorder should be defined and a description of how people define what is and is not disordered. I offer a potential solution to this problem by suggesting that the concepts of disorder and dysfunction are utilitarian constructs that arise out of the applied side nature of the health sciences.

The second major problem with the HDA is that it only provides a framework for defining and identifying biological disorders. For Wakefield, psychological disorders are of the same natural kind as biological disorders and both are adequately captured by the HDA. I argue this is a fundamental conceptual error. Psychology is no more fully reducible to biological theory than biology is fully reducible to chemical theory. As such, a definition of mental disorder that only captures biological dysfunctions is inadequate and misses an entire class of conditions, namely diagnosable mental disorders that are psychological in nature and that cannot be reduced to biological dysfunctions. Thus, whereas Wakefield (1992a, 1999a) equates disorder with disease, I argue for a differentiation between those mental disorders that result from broken biology (i.e., mental diseases) and those mental disorders that do not. This perspective offers a potentially powerful new way for conceptualizing mental disorders in general and may provide the beginning of a solution for how to define and differentiate psychiatric and psychological disorders. Before exploring these issues, however, we must turn our attention to the nature of the disorder construct.

The Nature of the Disorder Construct

Dysfunction is Not an Essentialistic Concept

Wakefield adopts a classical view of the disorder concept, summing up his position in his concluding paragraph in the JAB special issue: “Natural function refers to naturally selected effects, a concept well anchored in scientific theory, so dysfunction and disorder also refer to real phenomena.” (Wakefield, 1999b, p. 472). In concluding his target article, Wakefield (1999a) spelled out his claim that dysfunction is an essentialistic concept:

Concept theorists sometimes speak of underlying, theoretical-explanatory causal processes that unite a category as essences, and concepts defined in terms of such essential processes as essentialist concepts. The HD analysis claims that natural selection underlies natural functions and thus is crucial to attributions of dysfunction, as well; to this extent disorder is an essentialist construct. (However, disorder is not a purely essentialistic concept due to the harm component). (p. 397)

Wakefield is arguing that because evolutionary functionalism is an essentialist concept, then the dysfunction of evolved mechanisms must also be an essentialist concept. However, careful examination reveals this assertion to be problematic.

The reason this assertion is problematic is because the underlying causal structure of natural functions is different than the causal structure underlying the dysfunctions of those mechanisms. As I am sure Wakefield would agree, evolution through natural selection is a theory about the causal process that underlies biological organization, but it is not a theory about the infinite variety of causal processes that result in the breakdown of that organization. In fact, Wakefield (1999c) makes just this point in a recent article:

The theory of normal heart function and the account of heart disorders have very different logical structures because the domains of data have very different intrinsic properties. In the domain of normal functions, many diverse features are

likely to be understandable in terms of one elegant functional theory based on evolutionary design; in the domain of dysfunctions, each type of failure may become the subject of a theory, or many theories, all on its own. [italics added] (p. 970)

Thus, many diverse phenomena cause the dysfunctions of evolved mechanisms. This is important because, as Wakefield states, essentialistic concepts are labels for categories that are linked by underlying causal processes. Yet, breakdowns in evolved functional design are not essentialistic concepts as there is clearly not a global underlying causal process that accounts for or unites this set of conditions. Instead, dysfunctions, as defined by Wakefield, are merely linked to an essentialistic concept, natural function. Linking a construct to an essentialist concept and “discovering” an essentialist concept are two fundamentally different things.

If dysfunction is not an essentialistic concept, then can disorder be considered an essentialistic concept? Wakefield’s argument is that disorders are harmful manifestations of dysfunctions of an evolved mechanism. Thus one could, in theory, argue that disorders are caused by dysfunctions. If disorders are always caused by harmful breakdowns of evolved mechanisms, then perhaps disorders could be thought of as being essentialistic, even if dysfunctions are not. The problem here is that this argument is circular. If disorders are, by definition, harmful manifestations of bio-dysfunctions, then disorder is an essentialistic concept simply because of the definition of the concept itself. If the concept is only essentialistic because of its definition, then it loses its meaning. This analysis raises the question as to whether the HDA is intended to provide a description of the disorder construct or whether the HDA is a prescription for how the medical profession should define the disorder concept. It is to this question that we now turn.

Description or Prescription?: Wakefield's Unintentional Sleight of Hand and Other Conceptual Problems with the HDA

Ultimately, it is unclear if Wakefield (1999a, 1999b) was arguing that the HDA was supposed to describe when and how people make disorder attributions or was instead intended to be a prescription for how disorder attributions should be made. On several occasions, Wakefield (1999a) clearly implies that the HDA provides a framework for describing the types of conditions people believe constitute a disorder. Wakefield (1999a) bolsters this position with an advertisement of falsifiability: "I argue here that failure of a naturally selected function is necessary for disorder. This is a highly risky claim: it can be falsified by just one clear example of a disorder that is not an evolutionary dysfunction" (p. 376). Most of the reviewers interpreted the current version of the HDA as describing the components that individuals use when making disorder attributions. The fact that the major critique article by Lilienfeld and Marino (1999) offered an alternative Roschian analysis, which is clearly not a prescription for how to define disorder but instead is an attempt to describe how people make disorder attributions, lends credence to this interpretation. Yet, all of this is confusing because the original HDA proposition (Wakefield 1992a, 1992b) seemed to be a clear prescription for how the health sciences should define disorder.

Other reviewers noted this inconsistency. In fact, Sadler's (1999) critique was that the HDA went from a tool for discriminating disorder from nondisorder in its earlier prescription form (e.g., Wakefield, 1992a) to currently being a description of how disorder attributions are made. Wakefield's (1999b) response to Sadler, however, belies his intention on continuing to use the HDA as a prescription for how disorder attributions should be made: "Contrary to Sadler's concern that the HD analysis has no prescriptive bearing on the DSM-IV this discussion is very

much about eliminating false positives from DSM-V” (p.469). Indeed, it is clear from his writings that Wakefield believes that there are mental disorders that are consensually agreed upon [by definition in terms of their inclusion in the Diagnostic and Statistical Manual-IV (American Psychiatric Association, 1994)] which are, according to the HDA, are not mental disorders at all. In fact, Wakefield (1992b, 1997b, 1997c) has previously devoted several articles to just this point (see below for further discussion of this issue). Yet by stating that there are many DSM-IV diagnoses that do not contain design failures, Wakefield is demonstrating that the HDA fails as a description of how mental disorder attributions are currently being made. This is a serious self-contradiction.

It is important to note that in the Ground Rules section of the lead article, Wakefield (1999a) did insert the caveat that not every agreed upon disorder is, in actuality, a disorder and gave the example of masturbation being considered a disorder in Victorian times. Although this caveat appears reasonable at first glance, closer examination reveals that it gives Wakefield a loophole for deciding what is a real disorder and what is not. This, in turn, results in his argument being tautological. When examining disorders that surely involve a design failure and are seen as harmful (e.g., heart attack, broken bones), he can say his model is supported. On the other hand, when confronted with consensually agreed upon disorders that do not contain design failures (what Wakefield refers to as “false positives” in the DSM-IV), he argues that these conditions should not be considered disorders. If only disorders that have design failures are counted as real disorders, then the HDA is impossible to falsify as it creates a “heads I win, tails you lose” scenario. This is Wakefield’s sleight of hand. He advertises that his system is easily falsified, when in fact it is extremely difficult to falsify because he uses the HDA as either a

description of the types of phenomena people label as disordered or a prescription for how to make disorder attributions.

Using the HDA as both a description and a prescription raises another conceptual problem. When used as a description, the HDA is a proposal to describe how people make disorder attributions; when presented as prescription, it is a tool for making disorder attributions. The former is an explanation of a human social-cognitive process. The latter is a formula that supposedly represents something in the real world that guides professionals in their decision-making. These two phenomena, however, are quite separate entities and it is a violation of logic to suggest one is simultaneously describing how individuals are doing something and prescribing how individuals should be doing something.

To clarify, a valid theory describing disorder attributions, if there is such a thing, should explain why people make disorder attributions when they do, regardless of current conceptions as to whether or not the disorder attribution is seen as valid. Such a theory must operationalize what a disorder attribution is (e.g., an individual or group of people labeling a condition as something problematic or broken with the mind or body) and then explain when and why such attributions are made (e.g., the label justifies certain types of social reactions such as treatment or help, stigma, and/or control). To be successful, such an analysis would have to be applicable to all obvious cases of disorder attribution. For example, as Wakefield (1999a) notes, frequent masturbation was clearly seen as a disorder in Victorian times. A good theory of disorder attribution must be able to explain why this attribution was made, as well as why cancers, heart attacks and strokes are also labeled disorders. Lilienfeld and Marino's (1995; 1999) Roschian analysis is clearly an attempt to describe how individuals make disorder attributions.

A prescription for discriminating disorder from nondisorder is different. Such a tool identifies masturbation as conceptually different from cancers, heart attacks, and strokes, and dictates these phenomena be categorized accordingly. In this light, it becomes clear that the HDA is not really a description of how people make disorder attributions, but instead is a prescription for making distinctions between disorder and nondisorder.

The problem of vacillating between prescription and description did not go fully unnoticed. Reviewers Lilienfeld and Marino (1999) and Kirmayer and Young (1999) also noted that Wakefield is sometimes being descriptive and other times being prescriptive and that this disparity raises conceptual problems. I am revisiting these issues because the critique was raised only in equivocal terms, whereas I am claiming that vacillating between description and prescription is not a minor problem, but instead represents a fatal conceptual flaw that must be remedied if the analysis is to have merit. I am also suggesting that a key element of confusion in the debates between Wakefield, on the one hand, and Lilienfeld and Marino, on the other, is that their respective analyses represent an attempt to solve two fundamentally different types of problems. To remedy the situation, we must determine what type of construct the disorder concept is.

Disorder as a Utilitarian Construct

I argued earlier that even if one accepted Wakefield's definitions, dysfunction is not an essentialistic concept, but instead is simply linked to an essentialistic concept, natural function. I also stated that there is a fundamental difference between linking a construct to an essentialistic concept and discovering an essentialistic concept, although I did not elaborate. The fundamental difference can be highlighted by what purpose the concept serves. Pure essentialistic concepts like natural selection provide deep causal-explanatory frameworks for observational data. They

are pure theoretical constructs in that their utility exists in the degree of accuracy with which the concepts are able to account for observational data. Natural selection is a pure scientific construct because it functions to provide an algorithmic representation of change processes that can be tested for accuracy (i.e., organismic complexity should be a function of ancestral inclusive fitness).

In contrast to the concept of natural function, which arises out of the theoretical science of evolutionary biology, the pressure to define what is and what is not dysfunctional comes from the health sciences, which are applied sciences. The key difference between a theoretical science and an applied science is that the theoretical scientist uses the scientific method to describe change, whereas the applied scientist uses the scientific method to effect change. For the theoretician the accurate description of change is the end, whereas for the engineer the accurate description of change is a means toward some end. The health sciences are applied sciences because the goal of the health sciences is not the mathematical description of living and mental change processes, but instead is the application of knowledge in the service of increasing human biological, psychological, and social well being.

It is because of this goal that the health scientist is compelled to search for causes of detriments to functioning. To do this, the health scientist needs to be able to identify dysfunction. As the HDA points out, this can be done for biological disorders by linking dysfunction to the pure scientific concept of natural function. Such linkage is extremely useful because the ability to describe effects in scientific terms increases understanding and thus control. Yet such a boundary between function and dysfunction is not necessary from a purely theoretical/descriptive point of view. Due to the impossibility of the observed complexity arising by chance, theoretical biologists need to be able to account for natural function. However, this

does not mean that theoretical biologists need to account for the dysfunction of natural functions. That natural functions breakdown is not a mystery in the purely theoretical sense, but can be thought of as an inevitable consequence of entropy. Phenomena such as torn ligaments, blindness, and kidney failure are conceptually linked because of both natural selection's causal role in building ligaments, eyes, and kidneys and because of the goal of the health sciences to identify problems in functioning and to maximize human well-being. Thus, contrary to Wakefield's assertions, the set of dysfunctional conditions is conceptually linked by the applied side nature of the health sciences and thus cannot be thought of as essential property of nature.

If disorder is not a pure theoretical construct, then what type of construct is it? I suggest it can be thought of as a utilitarian construct. Utilitarian constructs are generally agreed upon definitions that function to facilitate problem definition, conceptual clarity, and communication among individuals. Such constructs can be thought of as definitional tools that assist a profession toward some applied goal. In this light, we can ask whether the HDA in its prescription form, serves a useful purpose for the health sciences.

It is clear that Wakefield does not view the disorder concept in these terms, although he does, of course, frequently state that there would be significant benefits to adopting his definition. In fact, and in direct contrast to the utilitarian approach offered here, Wakefield states in his opening comments that "It should ...be cautioned that the status of a condition as disordered or nondisordered from the HD or any other perspective has no necessary implication for the priority the condition deserves with respect to treatment, prevention, or policy" (Wakefield, 1999a, p. 374). I find this to be a remarkable caveat. Although the caveat is understandable in that there is not necessarily a one-to-one relationship between a condition being disordered and specific implications for treatment or policy, this does not seem to be

Wakefield's point. Instead, he wants to absolve himself from any political implications his analysis may have. Perhaps this is justified if one views dysfunction as a purely scientific construct. Yet, if the concept of dysfunction is a function of the applied side nature of the health sciences, then what does and does not constitute a disorder inevitably does and should carry some socio-political implications. Indeed, from the utilitarian perspective offered here, the definition of disorder is necessary precisely because it carries implications for treatment, prevention, and policy. With disorder defined as a utilitarian construct, we can now examine how well the HDA functions as a tool that is useful in problem definition and conceptual clarity.

The HDA as a Useful Definition for the Disease Construct

For Wakefield, disease is essentially synonymous with disorder (e.g., Wakefield, 1992a; 1999a). The two are synonymous for him because he views psychiatry and all the mental health sciences, including clinical psychology, to be extensions of medicine. Wakefield's (1997b) position is that there is a "critical need for the mental health professions to be seen as genuinely medical disciplines" (p. 643) and a main benefit of his definitional system is that it preserves that status. In line with this thinking, he states that a primary task of the DSM-IV is to "distinguish true mental disorders in the medical sense from the vast array of problematic but nondisordered human conditions often referred to as 'problems in living'" (Wakefield, 1997b, p. 634). As I discuss later, I strongly disagree with Wakefield's assessment that clinical psychology is an extension of medicine. However, before that issue is explored, it is useful to examine the HDA's utility in defining disorder in medicine, apart from psychiatry. For clarity, I will use the term disease to refer to medical disorders apart from mental disorders.

As indicated in the introduction, the HDA has much to offer. Those familiar with his work know that Wakefield offers many penetrating insights and has offered excellent critiques of

previous attempts at defining disorder and disease (e.g., Wakefield, 1992a, 1992b, 1993). When defining diseases, the HDA becomes an extremely powerful tool.

In accordance with the utilitarian view of the construct of disease, some criteria for what such a definition might accomplish should be established in order to evaluate the effectiveness of the HDA as a prescription for the definition of disease in the traditional medical sense. First, to be useful, a definition of disease should explain why certain conditions are currently classified as diseases, whereas other conditions are not. Another potentially useful function of a broad conceptual definition of disease is that it should provide a framework for understanding how the medical professions are currently organized. The definition should also capture why there are cross-cultural similarities and differences in what is defined as disease or illness. Also, although there will inevitably be fuzzy boundaries due to the nature of the disease concept, the definition should be effective at framing the relevant issues. How does the HDA fare with these criteria?

As a prescription for how the concept of disease should be defined by the medical profession apart from psychiatry, the HDA is extremely appealing and sophisticated. Given its simplicity, the HDA is incredibly successful at differentiating those conditions that are considered diseases from those that are not. It immediately identifies why heart attacks, broken bones, infections, malaria, cancers, and influenza are diseases, and why coughing with the flu, swelling following injury, and being weaker than average are not.

In addition to successfully differentiating current diseases from nondiseases, the analysis also captures why diseases are classified in reference to the various functional organ systems such as diseases of the circulatory system, diseases of the immune system, bone diseases, gastrointestinal diseases, and diseases of the reproductive system. Likewise, it explains why many of the various medical specialties, such as orthopedics, cardiology, gynecology, endocrinology, and

urology, are organized around functional systems in the body. In short, the HDA does an excellent job of explaining the current organization of medical disciplines.

As mentioned above, natural selection operating on genetic combinations through time is the fundamental theorem of biological science. As such, the HDA anchors the concept of disease to the fundamental concept in biological theory and captures how medical science conceptualizes biological problems. This linkage of medicine to biological theory is absolutely essential because it allows for scientific investigations into the many causes and consequences of disease, something necessary for the accumulation of objective knowledge and understanding.

The definition of disorder offered by the HDA is also of high pedagogic and heuristic value. It is parsimonious and organizes a tremendous amount of information. It is also quite possible that the concept can guide medical research and classification in the future. For example, the analysis allows for the crucial distinction between the actual disease (the destruction of cells by the flu virus) and the organism's defenses against disease [e.g., coughing when one has the flu; see Nesse & Williams, (1994); Wakefield, (1999b)].

By including the harm component, the HDA appropriately identifies medicine as an applied discipline and identifies concepts like disease as serving the purpose of guiding social action and policy in the service of human betterment. As Cosmides and Tooby (1999) pointed out, a human's illness is the infecting agent's health. The harm component highlights this human bias, which is often overlooked in medical science.

The combined value-science approach allows for conceptual clarity around confusing issues, such as the subjective versus the objective components of the disease concept. For example, the HDA clarifies why there are cross-cultural similarities and differences in what is identified as a disease. Because different cultures have different value structures, the analysis

rightly predicts that there will be some cross-cultural variation concerning the harmfulness and perceived harmfulness of certain dysfunctions. On the other hand, because human cultures are built by humans who all have essentially the same evolutionary history, one would expect much cross-cultural agreement on major dysfunctions, such as heart failure, blindness, or broken bones.

There are, of course, issues that remain to be addressed. For example, who makes the evaluation of harm, the individual, friends and family, society at large, the doctor? How much harm is necessary? Is annoyance harm? How much deviation from an evolutionary functional prototype is required to qualify as a dysfunction? How does one discriminate normal variation from dysfunction, particularly in areas that are not well understood? As underscored by these questions, there are elements of the HDA that still need clarification. However, problems with fuzzy boundaries are an inevitable product of the applied side nature of the concept, rather than a specific problem with the HDA. In fact, the HDA allows these problems to be appropriately framed and addressed and can be used to guide medical researchers and practitioners, bio-ethicists, policy makers, and lawyers in making complicated decisions pertaining to these issues. Thus, the HDA appears to provide an excellent definition for disease in medicine apart from psychiatry, better than any current alternative. It is an important contribution and something for which Wakefield should be highly commended.

The HDA Fails as a Definition for All Mental Disorders

Having argued that the HDA has strong utility for defining disease in medicine, we can now examine how the HDA fares for mental disorders. In accordance with the manner in which we analyzed the HDA for disease, we can ask the following questions about the HDA for mental disorders: (1) Does the HDA appropriately categorize the major mental disorders as it does with

major medical disease entities like heart attacks and strokes? (2) Does the HDA capture the way the mental health sciences are organized as it does for medicine? (3) Does the HDA link the applied mental health disciplines to the deep conceptual frameworks of the theoretical sciences of mind and behavior with as it does by linking medicine with evolutionary biology? In addressing these questions, I believe it will become clear that the HDA fails to account for what is generally meant by mental disorder.

All Mental Disorders as Breakdowns of Naturally Selected Mental Mechanisms?

A basic criterion for the HDA from a utilitarian perspective is that the current mental disorders should correspond to some degree with current diagnostic categories. And correspondence does occur for some disorders. Autistic disorder provides a relatively clear example. Recent models of autism have characterized the disorder as a condition of “mindblindness” which refers to the notion that individuals with autistic disorder have a serious dysfunction in their ability to understand and interpret social phenomena (Baron-Cohen, 1995). The theory is grounded in evolutionary analyses and presupposes that humans have special mental mechanisms that are devoted to processing socially relevant information. The observed symptoms associated with autistic disorder are thought to be the consequence of dysfunctions in these mechanisms. As such, this model conforms nicely to the HDA¹.

Schizophrenia is another example that seems to be appropriately classified by the HDA. A variety of different models of schizophrenia have been proposed in which there is a failure of some kind of natural mental mechanism. For example, Crow (1997; 2000) proposed a model of schizophrenia that, consistent with the HDA, suggests that schizophrenia is intimately related to

¹ Recent meta-analytic work by Yirmiya, Erel, Shaked, and Solomonica-Levi (1998) found that children with mental retardation of an unspecified etiology show very similar deficits in theory of mind abilities as children with autistic disorder. Such findings cast some doubt on the validity of the argument that deficits in theory of mind is the

the human capacity for language and results when there is a breakdown in the inter-hemispheric transfer of symbolic information. Regardless of their actual validity, these two models of schizophrenia and autistic disorder provide clear examples that conform nicely to the framework provided by the HDA as they both provide explanatory frameworks that attempt to account for the observed symptoms in terms of broken mental mechanisms. If all models of mental disorders fit similarly, then the HDA would likely have significant value for defining mental disorders.

Unfortunately for Wakefield, many mental disorders do not conform to the HDA. Lilienfeld and Marino (1995, 1999) provide insightful analyses of how the HDA fails for the learning disorders (reading, writing, and calculation disorders). Reading and writing are too culturally novel to have been fashioned by evolutionary processes, thus the mechanisms that allow humans to read and write must have originally evolved to serve some other purpose. As such, reading and writing disorders cannot be the consequence of the failure of naturally functioning reading and writing mechanisms.

Wakefield (1997a) attempts to salvage the HDA for learning disorders by arguing that the reason some children fail to learn how to read with instruction is that there likely is “some (as of yet unknown) mechanism with some (as of yet unknown) function...is unable to perform its function” (p. 278). His explanation is not very convincing and his defense is quite tautological. By his definition, all real disorders are the consequence of design failures. Wakefield agrees that learning disorders are real disorders and then argues that therefore there must be design failures present, despite the fact that there could not be specifically evolved reading and writing mechanisms and the fact that the neuro-cognitive data on the cause of such learning difficulties

central feature of autistic disorder. Nonetheless, the formulation offered by Baron-Cohen (1995) is a good example of a conceptualization of a mental disorder that conforms to the HDA.

are unclear. Wakefield engages in similar reasoning on numerous occasions (Houts & Follette, 1998).

What Wakefield fails to acknowledge is that the neuro-cognitive arrangements that currently result in reading and writing difficulties likely had little or no ramifications for functioning in preliterate societies. Thus reading and writing difficulties may very well result from normal functional variation in the architecture of the human mind that happens to cause problems in current literate environments. If this turns out to be the case, then learning disorders would not be the result of dysfunctions of evolved mechanisms but would be the result of a mismatch between current environments and the evolved architecture of the human mind. Wakefield readily acknowledges that mismatches should not be considered disorders from the perspective of the HDA (e.g., Wakefield, 1999a; 1999b). However, because Wakefield views learning disorders as real disorders, he presumes that neuro-cognitive arrangement associated with learning disorders are dysfunctions rather than currently problematic functional variants. As we will see, there are many such ambiguous examples that Wakefield seems willing to interpret only in a manner consistent with his definitional system.

There are also more “serious” forms of psychopathology that do not conform well to the HDA. Bulimia nervosa is a disorder characterized by binge eating episodes in which the individual feels a lack of control around food, makes problematic compensatory attempts to avoid weight gain, and possesses a self-evaluation that is unduly influenced by weight and body shape (APA, 1994). Current models of the disorder suggest that cultural emphases on appearance in general and thinness in particular, set the stage for certain individuals (particularly individuals who are younger, female, Caucasian, and have low self-esteem or demonstrate affective lability) to attribute their interpersonal success, at least in part, to their weight and body shape (e.g.,

Henriques & Calhoun, 1999; Stice, 1994). This attribution results in dieting. The dieting is often extreme, which sends signals to the body that food is scarce, which in turn sets the stage for a heightened hunger motive and difficulty controlling impulses to consume rich foods. These impulses then lead to binge eating, which in turn create more pressures to compensate for weight gain (Polivy & Herman, 1993).

If the HDA were adopted for mental disorders, one must decide that either bulimia is not a mental disorder or provide some theoretical rationale for the presence of a design failure of a mental mechanism associated with the etiology of disorder. The first choice does not appear to be a legitimate option. If a condition like bulimia nervosa is not considered a disorder than the entire concept of mental disorder would be so radically changed as to be unrecognizable. As for the second option, although one could, of course, weave stories about why this or that mental mechanism was failing to perform its naturally evolved function, the current models of bulimia nervosa do not conceptualize the disorder in anything that could be remotely considered to be the failure of an evolved mental mechanism. Instead, bulimia nervosa is conceptualized as a complex mixture of cultural, interpersonal, intrapsychic, and physiological elements, none of which is a design failure per se, but all of which interact to result in a psycho-behavioral syndrome that is seen as disordered, harmful, and requires treatment.

The learning disorders and eating disorders are not the only conditions that do not conform well the HDA. In fact, as mentioned above, Wakefield (1997b) readily acknowledged that if the HDA were adopted and enforced by the editors of the DSM, many currently diagnosed mental disorders would be either eliminated or significantly changed in future editions. Substance Abuse Disorders and Adjustment Disorders provide two immediate examples of

whole classes of disorders that would not be included in subsequent DSMs. As Wakefield writes in regards to Substance Abuse Disorders:

It is remarkable that the DSM-IV allows arguments with one's spouse about alcohol and drug use to be sufficient by itself for being diagnosed with a Substance Abuse problem. If you drink or smoke marijuana, your spouse can now give you a mental disorder simply by arguing with you about it and can cure you by becoming more tolerant! (Wakefield, 1997b, p. 640).

Remarkably, Wakefield generated examples for mood disorders, personality disorders, child and adolescent disorders, substance dependence disorders, and anxiety disorders, in addition to the adjustment and substance abuse disorders, that allow for a diagnosis but clearly do not involve design failures. Of course, in so doing, he is essentially demonstrating that failures of evolved mental mechanisms do not currently provide the conceptual underpinning of many mental disorders. His solution to the "overinclusiveness" problem is to "contextualize" the diagnostic categories and only include conditions in which one can reasonably infer the presence of a design failure. A primary justification Wakefield gives for this contextualization is that individuals diagnosed based on these criteria do not really have medical disorders and thus must be eliminated if the mental health sciences are to remain logically consistent with medicine.

Wakefield utilized Major Depression as an example for contextualizing a disorder. A Major Depressive Episode is defined as the presence of five out of nine symptoms present for two weeks or more (APA, 1994). Although life circumstances are generally not considered in making the diagnosis of a Major Depressive Disorder, there is one exception to the rule. If an individual has experienced a death of a loved one within the past two months then the individual should be considered bereaved, which is not a mental disorder. Wakefield (1997b) rightfully

criticized the DSM-IV for inconsistency in having such a caveat and pointed out that there are many other life circumstances, such as losing one's job, traumatic events, getting divorced, severe social isolation, abusive relationships, and financial difficulties, that often lead to marked depressive symptoms (e.g., Brown & Harris, 1978). In accordance with the HDA, Wakefield proposed that depressive disorders should only be diagnosed when the depressive symptoms are far more severe than the current living situation would warrant and that the diagnosis of MDD should be excluded if "the sadness response is caused by a real loss that is proportional in magnitude to the intensity and duration of the sadness response" (Wakefield, 1997b, p. 647). After all, it is only under such circumstances that one could reasonably infer that there is a malfunction of the depressive mechanism.

What Wakefield failed to acknowledge in his solution for "contextualizing" the diagnosis of depression is that depressive symptoms are frequently intertwined with serious problems in living. As one who works with individuals who present to an urban emergency room following a suicide attempt, I am quite familiar with patients who present with severe depressive symptoms. Virtually all such individuals (> 90 %) meet the diagnostic criteria for some form of mood disorder. Yet, not coincidentally, virtually all such individuals have extremely severe problems in living. High levels of severe physical and sexual abuse, prolonged unemployment, drug abuse and dependence, impoverished living conditions, little or no financial resources, little or no social support, physical illnesses, incarceration, systematic racism, and homelessness characterize the living situations and life histories of many of these individuals. Do such living conditions warrant severe depressive symptoms? If no, why not? If yes, then according to the HDA, many, if not most, of the suicide attempters I treat do not have a depressive disorder. Of course, such a conclusion would radically change our entire notion of the concept of depression

in particular and mental disorder in general. Thus the point is made again that Wakefield's definition of disorder does not conform to how mental disorders are currently operationalized, but instead requires a radical shift in how such conditions are conceptualized.

Furthermore, in accordance with the utilitarian perspective of disorder taken here, it is essential to point out that a significant problem immediately arises from Wakefield's "cure" for the DSM: What does society do with all the individuals who now do not meet criteria for having a "real disorder"? Wakefield does not address this issue. But it must be acknowledged that contextualizing these disorders and removing all the "false positives" would carry significant implications for real world issues like insurance reimbursement and whether individuals who are no longer classified as medically disordered are deserving of the sick role or not. One could just imagine a situation in which a treating professional has to say something such as, "I am sorry, Mrs. Jones. Because your husband beats you and because you have so few financial resources, your depression is warranted and thus is not a disorder. As such, insurance will not cover any treatment. If you lived with a caring husband in a nice house and were still depressed, then you would have a disorder and we could help." Of course, Wakefield would likely object to this criticism because of his caveat that whether or not a condition is disordered has no implications for him regarding treatment or policy. The response to such an objection is that this caveat does not reflect the reality of the situation. Society pays to treat individuals that are disordered and pays for research into 'real' disorders. Because Wakefield believes that the only real disorders are biological disorders, his solution would likely create significant problems in such issues as reimbursement for treatment.

In sum, Adjustment Disorders, Substance Abuse Disorders, Learning Disorders, Eating Disorders, Mood Disorders, Anxiety Disorders, and Personality Disorders are simply too many

classes of mental disorders in which there are specific conditions that cannot be readily accounted for by the HDA. Although one could argue that we simply have not found the underlying defective mental mechanisms that result in these disorders, this is currently a fantasy. Furthermore, there does not appear to even be circumstantial evidence of specific design failures for many currently diagnosed mental disorders and many conceptualizations of mental disorders exist that do not contain the failure of an evolved mental mechanism. Moreover, Wakefield himself readily agrees that many individuals are currently diagnosed with mental disorders that do not involve design failures. Given the number of major mental disorders that fail to correspond to the HDA, one is forced to agree with critics Kirmayer and Young (1999) who argued that the HDA “does not correspond to how the term disorder is used in psychiatry nosology or in clinician’s everyday practice and ...does not cover the territory to which the term reasonably could be applied” (p. 446). In short, one must conclude that the HDA fails as a currently useful global definition of mental disorder.

Does the HDA Capture How the Mental Health Sciences Are Organized?

In reviewing the utility of the HDA for medicine, I pointed out that the HDA corresponds well to the broad domains in medicine. The various organ systems are generally understood in terms of their evolved functional design and the various medical domains generally correspond to those organ systems. For example, the circulatory system corresponds to cardiology, obstetrics and gynecology correspond to the female reproductive system, orthopedics corresponds to the structural system, and so on. As such, the HDA provides a broad conceptual framework that links breakdowns in the parts that make up these organ systems.

We can now ask whether the HDA corresponds well with how the mental health sciences are organized. The correspondence is poor. The mental health sciences, unlike medicine, are not

really conceptually organized at all. Biopsychiatry, Radical Behaviorism, Cognitive Neuroscience, Social-Cognitive-Personality Psychology, Social Constructivism, Feminism, and Psychoanalysis (among other perspectives) exist as an unintegrated set of competing ideologies, each of which attempt to provide a framework for understanding both normal and abnormal functioning of mind and behavior. It is true that each perspective assumes that the nervous system is the product of natural selection. However, it is not true that each perspective conceptualizes the set of conditions currently labeled as mental disorders as the product of breakdowns in evolved mechanisms. In line with the earlier analysis regarding the relevance of the HDA to medicine, the biopsychiatric view is most consistent with, if not identical to, the HDA (Klein, 1999; Spitzer, 1999). However, several of the above listed perspectives are starkly at odds with Wakefield's formulation (e.g., Follette, 1996; Thakker, Ward, & Strongman, 1999). The existence of the other perspectives suggests that a purely medical or biological view of the mind fails to offer a complete explanatory framework (i.e., there are few if any 'nonmedical' perspectives on the dysfunction of other organ systems, such as the digestive tract). We now turn to the deep theoretical structure of the HDA and argue that Wakefield's analysis misses an entire layer of complexity that must be taken into consideration in the analysis of mental disorders, namely the mental or psychological layer of complexity.

Natural Selection is the Unified Theory of Life, Not Mind

A primary problem with the HDA as it currently stands is that there is no theory of the mind present in the analysis. Wakefield repeatedly comments that natural selection is the only known process that can explain the existence of mental mechanisms and thus all disorders are understood as the breakdown of evolved mechanisms. Yet there are clear examples that should immediately give us pause. The human capacity to read provides one such example. It is an

observational fact that most humans are capable of reading. In order to read, humans must have neurally based mental mechanisms that allow for the effective processing of written symbolic information. Yet virtually all theorists, including Wakefield, agree that it is impossible that humans possess reading mental mechanisms fashioned by natural selection. Thus reading provides us with clear evidence that mental mechanisms can somehow “evolve” as a consequence of experience during the individual’s lifetime. This suggests that Wakefield’s conclusion that natural selection is the only known process that causes the existence of mental mechanisms is either misleading, incomplete, or wrong.

The essence of the argument here is that mind represents a fundamentally different level of complexity than life. As Dawkins (1999) convincingly demonstrated, “living matter introduces a whole new set of rungs on the ladder of complexity” (p. 113) via natural selection operating on genetic combinations through time. Likewise, there are good reasons to believe that mental matter introduces yet another ‘set of rungs on the ladder of complexity’ apart from life. The reason is that just as the complexity of living matter is associated with a process of variation, selection, and retention giving rise to the genetic layer of algorithmic information, one could argue a parallel situation exists for mental matter. Behaviorists have long argued that the evolution of behavioral complexity over the course of an animal’s lifetime can be understood as the process of variation, selection and retention of behavioral responses (e.g., Donahoe, 1999; Skinner, 1974). Additionally, the cognitive neuroscience paradigm suggests that the brain is a neuro-chemical information processing device that functions to generate neuro-chemical representations of organism-environment relationships (Crick, 1994; Gazzaniga, 1995). Thus, one can readily argue that just as the gene can be considered the fundamental unit of biological information systems that cannot be reduced to the sum of its chemical parts, the neural impulse

exchanged between neural nets can be thought of as the fundamental unit or digit of psychological information systems that cannot be reduced to the sum of its biological parts.

The combination of the CNP with Behaviorist principles such as the Law of Effect strongly suggest that the evolution of behavioral complexity observed over the course of an animal's development represents a fundamentally different hierarchical level of complexity that is not fully explained by natural selection operating on genetic combinations. Because of this, it is a fallacy to suggest that one could use evolutionary theory as the sole guide in determining the functionality or dysfunctionality of animal behavior. Critics Richters and Hinshaw (1999) make a similar point, commenting that "The evolutionary cornerstone of J. C. Wakefield's (1999a) harmful dysfunction thesis is a faulty assumption of comparability between mental and biological processes that overlooks the unique plasticity and openness of the brain's functional design." (p. 438). It is important to note that this is not a dualistic argument in the Cartesian sense of the term. However, it does suggest that we can think about psychology being separated from biology in much the same way that biology is separated from chemistry. Despite all the advances in physics and chemistry, physical and chemical theories can not fully explain nor predict the biological complexity observed in organisms. Likewise, biological theory (natural selection and genetics) cannot fully explain nor predict the observed complexity in the behavior of the animal as a whole.

What does all this mean for the HDA? Put quite simply, it means that because the HDA incorporates only natural selection as a causal force it does not provide a complete framework for understanding the psychological layer of complexity and thus contains a greedily reductionistic (see Dennett, 1995) error that renders the HDA unworkable for a global definition of mental disorder. That there are so many classes of currently diagnosable mental disorders that

are not readily captured by the HDA and that there are so many nonmedical perspectives on mind and behavior supports this point.

Differentiating Mental Disorders from Mental Diseases

At first glance, the argument that there are mental disorders that are not biological disorders seems to inevitably raise the specter of dualism. After all, it is well agreed that the mind is the product of the brain, as opposed to some other nonphysical or metaphysical force. Thus, at least in some sense, everything that is psychological is biological. If this is true, wouldn't all psychological disorders have to be biological disorders? This is Wakefield's (1999d) argument for why psychological disorders must be biological disorders. But before we dismiss this possibility out of hand, an examination of the HDA itself should give us pause. Although everything that is biological is also physical, Wakefield rightly argues that not all biological disorders can be thought of as physical disorders. In challenging Szasz's (1974) argument that disorders are present only when there are physical lesions, Wakefield (1992a) demonstrated that there are lesions that are not disorders and disorders that do not appear to contain lesions. Instead, he eloquently argued that the concept of biological dysfunction must be anchored to functional analyses linked to the underlying causal processes of the phenomena in question, namely natural selection.

A similar conceptual issue is present in arguing that there are psychological disorders that are not biological disorders. If, as argued above, there is an essential difference between life and mind as there is between life and matter, then causal processes associated in the development of psychological complexity should provide the framework for understanding psychological disorders. If this argument is legitimate, then rigid, maladaptive behavioral patterns can emerge that do not involve the dysfunction of naturally selected mental mechanisms, but instead are the

result of breakdowns in the processes that give rise to behavioral complexity (i.e., learning). Such problems would be considered psychological or behavioral disorders that could not be reduced to biological theory.

In addition to these theoretical elements, it is important to note that the DSM-IV's (APA, 1994) definition of mental disorder allows for such a distinction when defining mental disorders. It states that, "Whatever its original cause, [the mental disorder] must currently be considered a manifestation of a behavioral, psychological, or [italics added] biological dysfunction in the individual" (p. xxi-xxii). Thus the framers of the DSM do not view psychological and behavioral disorders as redundant with, or identical to, biological disorders.

The Nervous System as a Behavioral Investment System

Before clarifying precisely what might constitute a psychological disorder that does not involve the breakdown of an evolved mental mechanism, it is important to provide at least a general framework for understanding nervous system complexity, or mind. In accordance with evolutionary theory, it is assumed that the nervous system evolved to compute and coordinate the behavior of the animal-as-a-whole. Furthermore, in accordance with such fields as neuropsychology, evolutionary psychology, and behavioral genetics, it is assumed that there is a basic bio-psychological architecture to the human mind. The occipital lobe processes visual information, the amygdala is intimately involved in the generation of fear and other emotions, the left cortical layer is associated with rational linguistic information processing, and so on. I strongly agree with Wakefield that harmful dysfunctions in these specific mental mechanisms can be considered to be mental diseases.

Although there is a general architecture, there is also a substantial amount of plasticity built into the system. Instead of mechanisms that rigidly dictate behavioral outputs, higher

animals have capacities to adjust their behavioral outputs contingent on the environmental effects that prior behavioral patterns have produced (e.g., Skinner, 1974). That is, the nervous system in higher animals has the capacity to measure “good” behavioral investments and produce more of them and measure “bad” behavioral investments and produce fewer of them². This remarkable capacity allows animals to fine tune their behavioral responses to the current environmental niche, even if the niche is a novel one. The ability of humans to live in modern environments is a testament to the remarkable power of the human capacity to adapt during ontogeny.

Behavioral Disorders as Maladaptive, Self-Perpetuating Behavioral Strategies

Although many animals are able to effectively fine-tune their behavioral responses to the current environmental niche, many animals are not able to do so. Impoverished, dangerous, or novel environments, unusual sequences or substantial changes in the contingencies of reinforcement, and problems more complex than the individual can solve can all result in maladaptive behavior patterns that emerge without a dysfunction of an evolved mental mechanism.

Consider the bulimic cycle of dieting and purging described earlier in relation to the behavioral investment model described here. This model suggests that bulimia nervosa can be effectively conceptualized as the product of conflicting goal states that produce behavioral solutions that paradoxically result in greater problems. Assume for the moment that high social status represents a goal state that is an intricate component of the evolved architecture of the mind. Further assume that individuals reference their status via social comparison mechanisms,

² “Good” and “bad” are in quotes because I am not referring to human values here, but the value system that evolutionary processes will build into the nervous system. “Good” behavioral investments can be defined as behavioral expenditures of energy that effectively move the animal toward an animal-environment relationship that

as well as by comparisons with one's past social influence. If we agree that women in Western cultures receive signals that thinness is associated with high social status, then one can see how losing weight is associated with increasing status and approached, whereas gaining weight becomes associated with decreasing status and avoided.

Of course, social status is not the only goal humans have. Food and weight set points likely represent even more fundamental motivators. Thus dieting, which serves to move the individual toward thinness and status, moves the individual away from weight-set points. The farther the individual moves away from weight and blood sugar set points, the greater the motive becomes to reduce this discrepancy. Controlling this impulse is especially difficult because the individual is bombarded with rich food cues. It is because of this sequence of events and conflicting motives that many theorists have argued that, paradoxically, dieting sets the stage for bingeing (e.g., Polivy & Herman, 1993). After the individual binges, the hunger motive is temporarily satisfied and the concerns about thinness and social status become salient. The individual will then often feel fat, ineffective, and unattractive (Stice, 1994). This leads to purging, more restrictive dieting, or the use of laxatives, which of course sets the stage for further bingeing. As the cycle repeats itself, it frequently has the ultimate effect of increasing the individuals' isolation and disconnection from others because they feel embarrassed and ashamed about their problematic eating patterns.

Bulimia is not the only disorder that can be understood as a dysregulation of the behavioral investment system. In a previous article (Henriques, 2000), I have argued that depressive reactions represent a behavioral shutdown response following repeated ineffectiveness, loss, or inescapable threat. I further argued that one can readily conceptualize

positively covaried with ancestral inclusive fitness, whereas "bad" behavioral investments are expenditures of

many depressive disorders as vicious cycles that result when the shutdown causes further loss which in turn leads to even greater shutdown. In such a formulation the shutdown mechanism is not malfunctioning (i.e., it is not responding when the individual is behaving effectively nor strongly responding in response to minor ineffectiveness), although it is producing a problematic, self-perpetuating cycle. Substance abuse and dependence, safety behaviors associated with anxiety disorders, and many aspects of personality disorders can be conceptualized as self-perpetuating, maladaptive behavioral cycles. Many behavioral (e.g., Fruzzetti, 1996), cognitive (e.g., Beck, 1976), and psychodynamic theorists (e.g., Wachtel & McKinney, 1992) have conceptualized psychological and relational problems in terms of vicious cycles of thoughts, feelings, and behavioral responses.

This model of the dysregulation of the behavioral system is characterized as a positive feedback loop in cybernetic terminology. The behavioral solutions adopted to achieve the initial goal ultimately have the paradoxical effect of creating more problems, which in turn only strengthens the problematic behavioral response. In the bulimic example, the lowering of status associated with the shame of purging and social isolation only serves to strengthen the focus on thinness as a way of increasing status, thus reinforcing the problematic cycle. Interestingly, in his article supporting Wakefield, Klein (1999) noted that many biological disorders can be conceptualized in the cybernetic terms of positive feedback loops. He wrote:

A dysfunction of the normal stabilizing negative feedback circuits causes the organism to either oscillate, as it is unable to find a stable set point, or reverse entirely into a pathological positive feedback cycle with a resultant wild excursion in either direction. (p. 422)

Thus, some behavioral disorders may share certain fundamental similarities with biological disorders without necessarily reflecting breakdowns in evolved design. This is important because it suggests that there may be ways to objectively define such behavioral disorders and differentiate them from deviant behaviors that would generally not be considered disordered.

It is important to state that, unlike with other biological systems, such maladaptive behavioral patterns are not design failures, do not include design failures, and cannot be understood as such. One may make a futile attempt to hold true to the HDA and argue that these behavioral patterns mean that the behavioral system is not functioning as it was designed to function. For example, one might suggest the appetite-regulatory system was not designed to purge perfectly good food, and thus is a dysfunction of an evolved mechanism. Or one might suggest that because the depressive response mechanisms evolved to prevent loss, the fact that such responses actually produce loss in certain circumstances are cases of design failures. These arguments, which Wakefield comes dangerously close to making on numerous occasions, are seriously flawed. Such arguments confuse a mechanism or system failing to produce “naturally intended effects” (which is a completely unworkable teleological position), with a mechanism failing to function in accordance with the complex functional design fashioned by evolutionary processes. The mental mechanisms described in the models of bulimia and depression offered here are theorized to be functioning as fashioned by evolutionary processes. They are simply producing effects that we judge to be harmful and could be considered ‘problems in living’ using Wakefield’s terminology. In order to remain consistent with the HDA, one must either demonstrate that these models are theoretically implausible or that such instances of vicious cycles resulting despite intact mental mechanisms should not be classified as genuine disorders.

Based on Wakefield's broad conception of what constitutes mental dysfunctions, it is possible that he would opt for a third choice and attempt to argue that bulimia and problematic depressive responses are failures of the behavioral system to function as designed. Perhaps, as he has argued before (Wakefield, 1997b; 1999c), he would suggest that these are 'software' problems or something similar, which he construes as design failures. However, this third option should be strongly rejected. Despite Wakefield's frequent intonations to the contrary, the neuro-behavioral system does not come equipped with a set of genetically preprogrammed behavioral responses to act in every possible situation. If this were the case, then the failure to produce such genetically instantiated preprogrammed responses would be design failures. But it is clear that this is not how the neuro-behavioral system is, or could be, designed. Instead, the neuro-behavioral system has been designed by evolutionary processes to be designed by experience (e.g., Dawkins, 1989; Skinner, 1974). Because of this additional layer of complexity, there is no evolved functional prototype for behaving a certain way. Thus, it is a fallacious oversimplification to state that because an individual is consistently behaving inappropriately or maladaptively that we can assume there is a dysfunction of an evolved mental mechanism. Instead, such maladaptive behavioral patterns are frequently the consequence of design and can be thought of as the inevitable price to be paid for behavioral flexibility. The ultimate point here is that, contrary to Wakefield's assertions, one can readily develop psychological models of disordered conditions that do not necessarily involve broken mental mechanisms and are not reducible to biological theory.

Psychiatry, Clinical Psychology, and Wakefield's Misdiagnosis of the DSM

It is Wakefield's political position that I most strongly disagree with. With little or no justification, he asserts that clinical psychology and social work are medical disciplines

(Wakefield, 1997b). He proceeds to argue that, because medical disciplines must effectively differentiate disordered from nondisordered conditions, the DSM and all mental health professions should adopt the HDA to be logically consistent with medicine at large. It is through framing the issues this way that psychiatry is placed as the supraordinate discipline and the bio-medical approach becomes the general framework with which all subdisciplines must be consistent.

In accordance with this socio-political starting point, Wakefield argued that characterizing mental disorders as something other than bio-medical problems would be a “radical solution” that would challenge psychiatry’s status as a medical discipline. Psychiatry may be reluctant to give up its status as a medical discipline, but as a clinical psychologist, it hardly feels like a “radical solution” to me. I was not trained as a medical doctor, I have never identified myself as a medical doctor, and I do not feel competent to treat medical problems. Thus, although Wakefield’s conception of the mental health field might result in him experiencing dissonance about treating nonmedical problems in living, it results in no dissonance in me. Instead, treating nonmedical psychological problems in living seems to be a very apt description for what I do.

I am not alone in this view and it is notable that Wakefield pronounces that clinical psychology is a medical discipline and essentially ignored all the reservations, complaints, and even hostilities that have been frequently expressed by clinical psychologists regarding the inadequacies and inappropriateness of a traditional medical approach for many important problems that require treatment. Humanists and psychotherapy integrationists (e.g., Wampold, 2001), psychodynamic theorists (e.g., Strupp, 1982), feminists (e.g., Liburd & Rothblum, 1995), and behaviorists (e.g., Follette, 1996) have all expressed significant concerns about the medical

approach³ applied to psychological problems. Renowned community psychologist George Albee (2000) recently went so far as to state that the profession of clinical psychology “sold its soul to the devil” (p. 248) when, as a consequence of socio-political and practical realities, the profession adopted the DSM nomenclature as its definitional system.

Part of Wakefield’s justification for all mental health professions adopting the HDA is that he believes that the concept of broken mental mechanisms is consistent with virtually all psychological positions, from behavioral to psychodynamic perspectives (e.g., Wakefield, 1992a). The problem here (in addition to the difficulties discussed earlier) is that neither behaviorists nor psychodynamic theorists tend to view psychopathology in terms of broken biological mechanisms. In his groundbreaking work in building previously unseen bridges between behaviorists and psychodynamic practitioners, Paul Wachtel (1997) noted that the rejection of the mental disease model is something both perspectives share. He points out that behaviorists often incorrectly criticize psychodynamic theorists for adopting a medical approach and argued that the connection between psychoanalysis and the medical profession is not based on a rational, close connection in thinking, but rather on sociological and economic factors. He further pointed out that superficial similarity in terminology between medicine and psychoanalysis does not mean shared conceptual similarity. For example, in reference to the term ‘symptom’, he wrote:

As it is used today by psychodynamic thinkers, the concept of symptom in no way implies a “disease.” The term refers to behaviors whose functional determinants are thought to include various aspects of the person as an ongoing psychological system, and whose change is therefore seen as often requiring a change in the

³ Although there are many different aspects of the medical approach, I use the term here in connection with

system...This is not a medical model, it is a psychological one. (Wachtel, 1997, p. 123-124).

The point here is that, despite Wakefield's (1992a; 1999c) claim that the notion of broken mental mechanisms derived from biological theory is essential for all conceptions of psychological disorder, neither behaviorists nor psychodynamic theorists tend to subscribe to a 'broken biology' framework for many if not most mental health problems.

In defending aspects of the DSM diagnostic system, Wakefield intonates that the idea that all genuine mental disorders reflect dysfunctions of evolved mental mechanisms is a "theory-neutral" stance (Wakefield, 1999c). Again his logic is that because all mental health disciplines are subdisciplines of psychiatry and thus are medical disciplines, and medical disciplines must treat genuine biological disorders, then, by definition, the only real mental disorders are biological disorders (Wakefield, 1999d). Unlike Wakefield, I do not see these as obvious truths. Instead, Wakefield and his supporters (e.g., Klein, 1999; Spitzer, 1999) are arguing for a theory-laden, biopsychiatric view of mental disorders. Furthermore, we should not be confused by Wakefield's scientific-sounding arguments regarding the nature of disorder. It is because of his political starting point, rather than any theoretical, epistemological or empirical arguments, that Wakefield believes that his conception of disorder is "theory-neutral" rather than theory laden.

Ultimately, Wakefield (1997b) misdiagnosed the problem underlying the DSM. The reason that the DSM is "overinclusive" is not "derived from a conflict within the views of the one person who more than anyone else influenced the conceptual structure of the DSM, Robert Spitzer" (Wakefield, 1997a, p. 643). With all due respect to Robert Spitzer, the issues are bigger

Wakefield's analysis, that the medical approach assumes that psychopathology is ultimately the result of

than he is. Instead, the DSM is “overinclusive” and thus inconsistent with a pure medical model because there are many conditions that are seen as disordered by professionals and the public alike, but are not true medical disorders in that they cannot be understood in terms of broken biology. Rather, when the behavioral investments that people make consistently fail to produce adequate solutions, but instead create additional problems in living and high levels of distress, then the individual is caught in a vicious positive feedback loop which can be considered to be a psychological or behavioral disorder. Such disorders justify diagnoses, treatment, and the sick role, but are not fully reducible to biological theory or describable in terms of malfunctioning evolved mental mechanisms.

It is true that up until this point clinical psychology has failed to provide a useful, coherent definition of mental disorder that can be widely adopted. Undoubtedly, this stems from the science of psychology lacking a well-defined subject matter and a coherent, unified theory of the mind. However, that science of psychology currently lacks a unified theory of the mind does not mean we should simply adopt the unified theory of life as a substitute guide in the meantime.

Conclusion

This paper has identified some important conceptual problems associated with the HDA. Dysfunction is not an essentialist concept and the HDA can only be considered as a prescription for distinguishing disorder from nondisorder, rather than as a description for when and how people make disorder attributions in general. These confusions result from Wakefield’s position that dysfunction is a pure scientific construct, rather than a utilitarian construct stemming from the applied side nature of the health sciences.

Taking a utilitarian view, the HDA does an excellent job of characterizing disorder in medical fields other than psychiatry. It provides a framework for why the medical professions are organized around various functional systems in the body and, by including the harm component, the HDA captures the human value system that underlies the applied side nature of medicine. Finally, by anchoring the concept of disease to the fundamental theory of life, medicine is appropriately linked to its pure science parent discipline, biology.

Although the HDA seems to be of high utility for defining disease in medicine, it fails as a global construct for mental disorders. Some mental disorders, such as autistic disorder and schizophrenia, seem to be readily captured by the HDA. But many major mental disorders, ranging from the substance use disorders to bulimia to depression to many personality disorders, cannot be readily understood as the consequence of design failures of mental mechanisms. The HDA also fails to capture how the mental health sciences are organized. I further argued that the HDA contains greedily reductionistic errors because it fails to incorporate a model of mental complexity and incorrectly implies that psychology is fully reducible to biological theory.

An alternative conceptualization of mental disorders was offered which suggested that many mental or behavioral disorders could be thought of as dysregulations of the behavioral investment system in which individuals develop maladaptive solutions to problems in their environment. Because of the open design of the neuro-behavioral system and the fact that there is not an evolved functional prototype for behaving in all situations, these maladaptive behavioral patterns cannot be thought of as deviations from evolved behavioral prototypes or neuro-behavioral mechanisms. It was also argued that, contrary to Wakefield's assertions, many, if not most, clinical psychologists do not implicitly or explicitly adopt a medical approach to all mental disorders. Finally, it was argued that it is because of Wakefield's political position

regarding the need of mental health professions to be seen as genuine medical disciplines that he incorrectly views his approach as theory neutral rather than theory laden and misdiagnoses the problem underlying the DSM.

It appears that it was no accident that the reviewers in the special issue who criticized the HDA tended to focus on how the analysis failed to capture the nature of mental disorders, whereas Wakefield tended to focus on more traditional medical problems to demonstrate its success and utility. From my perspective, the fact that Wakefield's analysis fails for many mental disorders does not mean that the HDA is useless but rather that Wakefield overextended the reach of the concept.

Using the HDA as a solution to the disease concept, but not accepting it as the definition for all mental disorders, opens up a very interesting possibility. By effectively defining the disease construct, those in psychiatry and psychology now have a tool for discriminating disease from nondisease. This analysis points out that although many mental disorders are diseases because they almost certainly involve breakdowns in the basic bio-psychological architecture of the human mind (e.g., schizophrenia and autistic disorder), many mental disorders are not diseases in the traditional medical sense (e.g., at a minimum, Wakefield's long list of false positives) because they do not involve such dysfunctions. Thus, the HDA offers a potential explanation for why psychiatry has always struggled with the question of whether or not mental disorders are conceptually equivalent to diseases in general medicine. The answer, ultimately provided by the HDA, is that some mental disorders are diseases, whereas some mental disorders are not.

References:

- Albee, G. W. (2000). The Boulder model's fatal flaw. American Psychologist, *55*, 247-248.
- American Psychiatric Association (1994). Diagnostic and statistical manual of mental disorders (4th ed.). Washington, DC: American Psychiatric Association.
- Baron-Cohen, S. (1995). Mindblindness: An essay on autism and theory of mind. Cambridge, MA: MIT Press.
- Beck, A. T. (1976). Cognitive therapy of the emotional disorders. New York: Meridian.
- Brown, G. W., & Harris, T. (1978). Social origins of depression: a study of psychiatric disorder in women. London: Tavistock Publications.
- Cosmides, L., & Tooby, J. (1999). Toward an evolutionary taxonomy of treatable conditions. Journal of Abnormal Psychology, *108*, 453-464.
- Crick, F. (1994). The astonishing hypothesis: The scientific search for the soul. New York: Simon & Schuster.
- Crow, T. J. (1997). Is schizophrenia the price the Homo Sapiens pay for language? Schizophrenia Research, *28*, 127-141.
- Crow, T. J. (2000). Schizophrenia as the price Homo sapiens pay for language: A resolution of the central paradox in the origin of species. Brain Research Reviews, *31*, 119-129.
- Dawkins, R. (1989). The selfish gene (2nd Edition). Oxford: Oxford University Press.
- Dawkins, R. (1999). The extended phenotype: The long reach of the gene (2nd Edition). New York: Oxford University Press.
- Dennett, D. (1995). Darwin's dangerous idea: Evolution and the meanings of life. New York: Simon & Schuster.
- Donahoe, J. W. (1999). Edward L. Thorndike: The selectionist connectionist. Journal of

the Experimental Analysis of Behavior, 72, 451-454.

Follette, W. C. (1996). Introduction to the special section on the development of theoretically coherent alternatives to the DSM system. Journal of Consulting and Clinical Psychology, 64, 1120-1132.

Fruzzetti, A. E. (1996). Causes and consequences: Individual distress in the context of couple interactions. Journal of Consulting & Clinical Psychology, 64, 1192-1201.

Gazzaniga, M. S. (1995). The cognitive neurosciences. Cambridge: MIT Press.

Ginsburg, B. E., & Carter, B. F. (1987). Premenstrual syndrome: Ethical and legal implications in a biomedical perspective. New York, NY, US: Plenum Press.

Henriques, G. H. (2000). Depression: Disease or behavioral shutdown mechanism? Scipolicy: The Journal of Science and Health Policy, 1, 152-165.

Henriques, G. R., & Calhoun, L. G. (1999). Gender and ethnic differences in the relationship between body and self-esteem. Journal of Psychology, 133, 357-368.

Hetherington, Marion M. (1993). In what way is eating disordered in the eating disorders? International Review of Psychiatry, 5, 33-49.

Houts, A. C., & Follette, W. C. (1998). Mentalism, mechanisms, and medical analogues: reply to Wakefield (1998). Journal of Consulting and Clinical Psychology, 64, 1169-1176.

Kendell, R. E. (1975). The concept of disease and its implications for psychiatry. British Journal of Psychiatry, 127, 305-315.

Kirmayer, L. J., & Young, A. (1999). Culture and context in the evolutionary concept of mental disorder. Journal of Abnormal Psychology, 108, 446-452.

Klein, D. F. (1999). Harmful dysfunction, disorder, disease, illness, and evolution. Journal of Abnormal Psychology, 108, 421-429.

- Koski, P. R., & Mangold, W. D. (1993). A sociological perspective: Antisocial behavior and violence. In A. S. Bellack & M. Hersen (Eds). Psychopathology in adulthood, (pp. 104-120). Needham Heights, MA: Allyn & Bacon.
- Liburd, R., & Rothblum, E. (1995). The medical model. In E. J. Rave & C. C. Larsen, (Eds.) Ethical decision making in therapy: Feminist perspectives. (pp. 177-201). New York: Guilford Press.
- Lilienfeld, S. O., & Marino, L. (1995). Mental disorder as a Rochian concept: A critique of Wakefield's "harmful dysfunction" analysis. Journal of Abnormal Psychology, *104*, 411-420.
- Lilienfeld, S. O., & Marino, L. (1999). Essentialism revisited: Evolutionary theory and the concept of a mental disorder. Journal of Abnormal Psychology, *108*, 400-411.
- Nesse, R. M., & Williams, G. C. (1994). Why we get sick: The new science of Darwinian medicine. New York: Random House.
- Polivy, J., & Herman, C. P. (1993). Etiology of binge eating: Psychological mechanisms. In C. G. Fairburn and G. T. Wilson (Eds.), Binge eating: Nature, treatment and assessment. (pp 173-205). New York: Guilford Press.
- Richters, J. E., & Hinshaw, S. P. (1999). The abduction of disorder in psychiatry. Journal of Abnormal Psychology, *108*, 438-445.
- Sadler, J. Z. (1999). Horsefeathers: A commentary on "Evolutionary versus prototype analyses of the concept of disorder." Journal of Abnormal Psychology, *108*, 433-437.
- Scadding, J. G. (1990). The semantic problem of psychiatry. Psychological Medicine, *20*, 243-248.
- Sedgwick, P. (1982). Psycho politics. New York: Harper & Row.

- Skinner, B. F. (1974). About behaviorism. New York, NY: Alfred A. Knopf.
- Spitzer, R. L. (1999). Harmful dysfunction and the DSM definition of mental disorder. Journal of Abnormal Psychology, *108*, 430-432.
- Stice, E. (1994). Review of the evidence for a sociocultural model of bulimia nervosa and an exploration of the mechanisms of action. Clinical Psychology Review, *14*, 633-661.
- Strupp, H. H. (1982). Is the medical model appropriate for psychoanalysis? Journal of the American Academy of Psychoanalysis, *10*, 123-128.
- Szasz, T. S. (1974). The myth of mental illness: Foundations of a theory of personal conduct (Rev. ed.). New York: Harper & Row.
- Thakker, J., Ward, T., & Strongman, K. T. (1999). Mental disorder and cross-cultural psychology: A constructivist perspective. Clinical Psychology Review, *19*, 843-847.
- Wachtel, P. (1997). Psychoanalysis, behavior therapy and the relational world. Washington, D.C.: American Psychological Association.
- Wachtel, P., & McKinney, M. K. (1992). Cyclical psychodynamics and integrative psychodynamic therapy. In J. C. Norcross & M. R. Goldfried (Eds.), Handbook of psychotherapy integration (pp. 335-372). 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 the 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. (1997a). When is development disordered? Developmental psychopathology and the harmful dysfunction analysis of mental disorder. Development and Psychopathology, 9, 269-290.
- Wakefield, J. C. (1997b). Diagnosing DSM-IV--Part I: DSM-IV and the concept of disorder. Behaviour Research and Therapy, 35, 633-650.
- Wakefield, J. C. (1999a). Evolutionary versus prototype analyses of the concept of disorder. Journal of Abnormal Psychology, 108, 374-399.
- Wakefield, J. C. (1999b). Mental disorder as a black box essentialist concept. Journal of Abnormal Psychology, 108, 465-472.
- Wakefield, J. C. (1999c). Philosophy of science and the progressiveness of the DSM's theory-neutral nosology: response to Follette and Houts, part I. Behaviour Research and Therapy, 37, 963-999.
- Wakefield, J. C. (1999d). Philosophy of science and the progressiveness of the DSM's theory-neutral nosology: response to Follette and Houts, part II. Behaviour Research and Therapy, 37, 1001-1027.
- Wampold, B E. (2001). The great psychotherapy debate: Models, methods, and findings. Mahwah, NJ: Lawrence Erlbaum Associates.
- Yirmiya, N., Erel, O., Shaked, M., & Solomonica-Levi, D. (1998). Meta-analyses comparing theory of mind abilities of individuals with autism, individuals with mental retardation, and normally developing individuals. Psychological Bulletin, 124, 283-307.