



## The Irrelevance of the Medical Model of Mental Illness to Law and Ethics

Aaron S. Greenberg\* and J. Michael Bailey\*\*

The United States Supreme Court in 1988 considered a challenge to the Veterans Administration's determination that in some cases alcoholism constitutes "willful misconduct."<sup>1</sup> Such willful misconduct would disqualify the alcoholic from an extension of time within which to receive certain veterans benefits. Although the Court determined that the case "does not require the Court to decide whether alcoholism is a disease whose course its victims cannot control,"<sup>2</sup> the case did renew, in both the popular and professional literature, the dispute concerning this and related issues.<sup>3</sup> The proposition that alcoholism is a disease, or that it has physiological correlates or a genetic basis, is thought to entail legal and social consequences, such as the conclusion that the alcoholic's conduct is not "willful," and that governmental benefits should not be denied him or her on that basis.

The alcoholism-as-disease dispute is one instance of a more general debate concerning all mental disorders.<sup>4</sup> Countless pages have been written about

---

\*Attorney.

\*\*Assistant Professor, Department of Psychology, Northwestern University, 2029 Sheridan Road, Evanston, Illinois, 60208-2710, USA.

Address correspondence to J. Michael Bailey.

<sup>1</sup>Traynor v. Turnage, 485 U.S. 535 (1988).

<sup>2</sup>*Id.* at 552. The actual issue before the Court (in addition to a jurisdictional question) was whether the Veterans Administration had the statutory authority to make such a determination and whether that determination violated a federal non-discrimination statute.

<sup>3</sup>See, e.g., Fingarette, *Alcoholism: The Mythical Disease*, THE PUBLIC INTEREST, Spring, 1988, at 3; Madsen, *Thin Thinking About Heavy Drinking*, THE PUBLIC INTEREST, Spring, 1989, at 112; Marwick, *Court Ruling Expected Soon in Alcoholism Case*, 2590 J. A.M.A. 1436 (1988); Seessel, *Beyond the Supreme Court Ruling on Alcoholism as Willful Misconduct: It Is Up to Congress to Act*, 259 J. A.M.A. 248 (1988); Wright, *Alcohol and Free Will*, THE NEW REPUBLIC, Dec. 14, 1987, at 14; Vatz and Weinberg, *Confusion Over Alcoholism: Psychiatry, Medicine, and the Law Disagree*, U.S.A. TODAY, Sept. 1989, at 68.

<sup>4</sup>Throughout this paper we refer to "mental disorders" for want of a better term. By this term we mean only to identify those behavior patterns commonly referred to as "mental illnesses" or "insanity." Our use of this term should not be taken to attribute any characteristics to such behavior patterns, and we specifically take no position on the question of whether such behavior patterns exhibit less order in any sense than other behavior patterns.

whether mental disorders are diseases, or whether the so-called medical model of mental disorders is appropriate.<sup>5</sup> The answer to that question is thought to imply various legal, social, and ethical consequences in areas such as legal and ethical non-responsibility and the propriety of involuntary commitment of the mentally disordered. Thomas Szasz, for example, has argued that mental disorders are not illnesses in any literal sense and has on that basis attacked the insanity defense and the practice of involuntary commitment.<sup>6</sup> Recent examples dealing with specific mental disorders include post traumatic stress disorder and the "battered wife" syndrome. The identification by psychiatric authorities of these behavior patterns as diseases or disorders, complete with lists of symptoms, is thought to support the conclusion that proof that a defendant's illegal act was caused by such a syndrome should relieve him or her of legal responsibility.<sup>7</sup>

It is our contention that notwithstanding this heated and protracted debate over the medical model of mental disorders, the propriety of classifying mental disorders as illnesses has no legitimate implications for questions of social and legal policy such as criminal responsibility and involuntary commitment.

First, we will attempt to show that, as a general matter, determining whether or not mental disorders are illnesses yields no new information about the characteristics of mental disorders or the proper treatment by society of those afflicted with them. The second part of the paper takes up the question of whether the central feature of the medical model, biological causation, has legitimate legal or societal implications. We conclude that the mere fact of biological causation of mental disorders has no such implications. Finally, we argue that while particular types of biological causation could, in principle, have legal or policy implications, the matter is much more complex and much less clear than generally believed.

### The Analysis of Mental Disorders as Illnesses

Are mental illnesses *illnesses* properly so-called? This can be understood as a linguistic question, namely: Is it proper to apply the word *illness* to a given mental disorder or, do mental disorders come within the proper definition of the word *illness*?

The first step in any determination of whether mental disorders are illnesses is to arrive at an acceptable definition of *illness*. This would entail specifying an exhaustive list of the defining characteristics of the concept of illness, what-

---

<sup>5</sup>See, e.g., N. ANDREASEN, *THE BROKEN BRAIN* (1984); R. KENDELL, *THE ROLE OF DIAGNOSIS IN PSYCHIATRY* (1975); T. SZASZ, *INSANITY: THE IDEA AND ITS CONSEQUENCES* (1987); Boorse, *What a Theory of Mental Health Should Be*, J. FOR THEORY SOC. BEHAV 49 (1975); KLEIN, *A PROPOSED DEFINITION OF MENTAL ILLNESS*, in *CRITICAL ISSUES IN PSYCHIATRIC DIAGNOSIS* 41 (1978); MOORE, *DISCUSSION OF THE SPITZER-ENDICOTT AND KLEIN PROPOSED DEFINITIONS OF MENTAL DISORDER (ILLNESS)*, *id.*, at 85; SPITZER AND ENDICOTT, *MEDICAL AND MENTAL DISORDER: PROPOSED DEFINITIONS AND CRITERIA*, *id.* at 15.

<sup>6</sup>E.g., T. SZASZ, *LAW, LIBERTY AND PSYCHIATRY* (1963).

<sup>7</sup>For a popular account of recent developments in this area, see Saletan and Waltzman, *Marcus Welby, J.D.*, *THE NEW REPUBLIC*, April 17, 1989, at 19.

ever those might be.<sup>8</sup> Once such a definition of *illness* is arrived at, we must next determine whether mental disorders come within that definition.

Thus, the analysis of whether mental disorders are illnesses typically involves two steps:

1. What are the characteristics for which *illness* stands?
2. Do mental disorders possess those characteristics?

For example, Szasz argues that: 1) An illness entails a physical lesion on some part of the body; and 2) Mental disorders do not cause or arise from such lesions (or at least this has not been demonstrated). Szasz, therefore, concludes that mental disorders are not *illnesses* properly so-called.<sup>9</sup> So to assert that mental disorders are illnesses is to say that all illnesses possess characteristics A, B, and C (again, this must be an exhaustive list of the defining characteristics), and that mental disorders possess these same characteristics A, B, and C.

### The "Illness" Analysis is Either Circular or Fallacious

Suppose that, under such an analysis, it were established that mental disorders are illnesses. From this proposition, we could legitimately deduce that mental disorders possess the defining characteristics of illness, A, B, and C (whatever those might be). In fact, this (together with any propositions implied by the conjunction of A, B, and C<sup>10</sup>) is *all* that we could deduce from our definition of mental disorders as illnesses. Could we deduce, for example, that medical doctors ought to treat mentally disordered people, or that dangerous mentally disordered people should be confined or treated against their will, or that mentally disordered people should not be held responsible, morally or legally, for their actions? Clearly we could not, *unless* those things are deducible from A, B, and C themselves. Those consequences would follow (if indeed they do) solely from the characteristics possessed by mental disorders and not, to any extent, from the fact that those characteristics happen to be the definitional characteristics of the concept of illness. If, for example, all illnesses ought to be treated by medical doctors, it is not because all illnesses are illnesses, but because all illnesses have some characteristic that makes them properly treatable by such doctors. Similarly, if ill people are not morally responsible for their illnesses and symptoms, it is not because what they have

---

<sup>8</sup>One or more of these characteristics could be negative, *e.g.*, one might be the characteristic of *not* possessing some other characteristic. Strictly speaking, "illness" could also be defined in the alternative, *e.g.*, an illness is something that is characterized by A, B, and C *or* by D, E, and F. If that were the case, the analysis that follows would simply be applied to each of the alternative definitions.

<sup>9</sup>*See, e.g.*, T. SZASZ, *supra* note 5.

<sup>10</sup>These implications from the defining characteristics of illnesses may be strictly logical (*e.g.*, if all illnesses are biologically caused, it cannot be the case that illnesses are not biologically caused). Alternatively, such an implication may be the consequence of an empirical hypothesis we may have about a characteristic. For example, we might maintain that one defining characteristic of illnesses is that they require medical treatment. We might also hold the empirical hypothesis that only physicians are capable of providing competent medical treatment. In that case, our characteristic of illnesses, together with our empirical hypothesis, would imply that only physicians can competently treat illnesses. In this paper, characteristics should be taken to include all such logical and empirical inferences which may be drawn from them.

are illnesses but because some characteristic of what they have absolves them of responsibility (*e.g.*, they have little or no control over their illnesses and symptoms).

Now, given this, suppose we want to answer the question of whether mentally disordered people are morally responsible for their actions and suppose further that we agree that moral responsibility turns on the characteristic of uncontrollability of symptoms, which we shall call C. The illness analysis might go as follows:

1. The defining characteristics of illness are A, B, and C (by hypothesis).
2. Mental disorders are characterized by A, B, and C (by hypothesis).
3. Mental disorders are illnesses (From 1 and 2).
4. People with characteristic C are not responsible (by hypothesis).
5. Mental disorders have characteristic C (from 1 and 3).
6. People with mental disorders are not responsible (from 4 and 5).

Obviously, this is an inefficient and, given our hypotheses, circular method of answering our original question about the responsibility of mentally disordered people. One need not be a skilled logician to arrive at the following superior alternative:

1. Mental disorders are characterized by A, B, and C (by hypothesis).
2. People with characteristic C are not responsible (by hypothesis).
3. People with mental disorders are not responsible (from 1 and 2).

Thus, the illness analysis adds several unnecessary steps to the enterprise of determining the legitimate ethical and legal consequences of mentally disordered behavior. Such a determination must follow from the characteristics of mental disorders; whether those characteristics qualify mental disorders as illnesses is irrelevant and unnecessarily introduces a very complex and controversial question, namely, the proper definition of *illness*.

The illness analysis, as sometimes applied, can be worse than irrelevant and circular; it can be fallacious and misleading. The 6-step illness argument outlined above is inefficient and unnecessarily complicated, but it at least is logically valid. Consider the illness advocate who argues as follows: Physical illnesses have characteristics A, B, C, and D (where D is the inability of the ill person to control the onset and symptoms of the illness, and is not implied, logically or empirically [*see note 10, supra*], by A, B, and C). Mental disorders have characteristics A, B, and C. They therefore share sufficient qualities with physical illnesses (our paradigms for illness) to legitimately be called illnesses. Since mental disorders are illnesses, like physical illnesses, these mental illnesses also possess characteristic D, which is possessed by all other illnesses. Thus, mentally ill people, like physically ill people, cannot control the onset or symptoms of their illnesses. In a more schematic form:

1. All illnesses possess A, B, C and D (by hypothesis).
2. Mental disorders possess A, B and C (by hypothesis).

3. Mental disorders are illnesses (?).
4. Mental disorders possess D (from 1 and 3).

Obviously, 1 and 2 above do not imply 3, and the illness advocate's argument is fallacious. One cannot validly argue that, because one has determined that mental disorders are illnesses, mental disorders possess a certain characteristic of all other illnesses, if that characteristic was not a member of (or implied by) the set of characteristics used to make the initial determination that mental disorders are illnesses. And, if that characteristic was a member of (or implied by) that set of defining characteristics, we are back to the inefficient and circular (though valid) illness analysis discussed above.

It would, therefore, appear that those who believe that the mental-disorder-as-illness debate is a substantively important one, are faced with a dilemma, for it must be the case that either:

1. The characteristic imputed to mental disorders by defining them as illnesses *was* used to determine that mental disorders are illnesses (or is implied by the characteristics so used), in which case our definition tells us nothing that we did not already know before making it; or
2. The characteristic imputed to mental disorders by defining them as illnesses *was not* used to determine that mental disorders are illnesses (and is not implied by the characteristics so used), in which case it simply does not follow from that definition that mental disorders possess the characteristic in question.

Thus, determining that mental disorders possess a given characteristic because we have defined them as illnesses is either circular or erroneous.

Those who use the illness analysis in an attempt to derive (or disprove) new information about mental disorders misconceive the role of classification. As shown above, classification in itself cannot generate new information about the entities classified. For any class, if we know that characteristics A and B qualify an entity as a member of the class and we know that all members of the class have characteristic C, we can recast our hypothesis into one that simply says that all entities with A and B have C. This is one variety of the circular case discussed above, in which the characteristics used to make the classification (A and B), together with a hypothesis about those characteristics (everything with A and B has C), in themselves (*i.e.*, without the classification) imply the new characteristic (C). The classification aspect is eliminable.

This is not to say that classification cannot be useful in certain kinds of cases. First, a general term can of course be useful in discourse as a label which those who use it stipulate will stand for specific instances, even if the classification gives us no new information. This, however, is not what illness analysts are seeking to do. This is evident from two facts about the illness debate. First, the participants believe that a determination that mental disorders are illnesses will establish some important, non-linguistic, fact about such disorders. Second, the illness analysts argue about what *illness* means or about what illnesses are, rather than simply agreeing on a label (any word would do) for the sole purpose of discourse.

A second useful aspect of classification (*e.g.*, of organisms by species, genus, etc., or of elements on the periodic table) is that it can provide a helpful way of organizing or conceptualizing large amounts of information in a systematic, comprehensible form. We would perhaps expect this use of classification to be of the greatest utility in cases where the entities being classified are naturally occurring kinds (such as organisms and elements), whose properties and relationships to one another are determined by or reflected in their physical structures, quite apart from how we think or talk about them. These cases may be contrasted with classifications based on conceptual constructs like *illness*, whose characteristics are determined in large part by stipulation or linguistic convention. In any case, at least some of these classifications can clarify or make easier to see connections among entities and characteristics that we already possessed the logical materials to make. What such classifications, including the illness analyses, cannot do is give us new information about such connections which could not have been determined before the classifications were made.

### Examples

By way of illustration of the foregoing analysis of the illness debate, consider the following examples. Blashfield states:

The importance of defining the concepts of *disease* and *mental disorder* is that their definition will help to resolve the *demarcation problem*. In other words, these definitions will help draw a boundary between persons who should be seen by mental health professionals and persons who should not.<sup>11</sup>

But can such definitions really help draw this boundary? If people with mental disorders should be seen by mental health professionals, it is due, let us say, to such factors as (a) the judgment that mental disorders are bad (*i.e.*, people should want to be rid of them), and (b) the belief that mental health professionals can effectively help people to be rid of them. Why spend our time wondering about the proper definition of *disease* and whether mental disorders are diseases? If people with diseases should be seen by health professionals because all diseases are bad and treatable (let us say), we can validly elongate our chain of reasoning to find that mental disorders are bad and treatable and therefore are diseases and that, because they have diseases, people with mental disorders, like people with any other disease, should be seen by health professionals (in this case, mental health professionals). But in so doing, we gain nothing and we add unnecessary levels of complexity (defining *disease* and determining whether mental disorders come within that definition) to our analysis. And, clearly, if we classify mental disorders as diseases for reasons not related to badness and treatability, the next step of finding that the mentally disordered should seek professional help because they have diseases is fallacious.

For a second example of the overcomplexity of the illness analysis, we turn

---

<sup>11</sup>R. BLASHFIELD, *THE CLASSIFICATION OF PSYCHOPATHOLOGY* 77 (1984).

to Szasz. One of Szasz's arguments against involuntary commitment (at least as stated in some of his writings) appears to be as follows:

1. Illnesses necessarily involve lesions in or on the body.
2. So-called mental illnesses do not involve such lesions.
3. Mental illnesses therefore are not real illnesses (they are, instead, illnesses in only a metaphorical or mythical sense).
4. Therefore, laws dealing with, for example, involuntary commitment should be repealed because they are premised on people having mental illnesses, which do not exist (other than metaphorically).<sup>12</sup>

If Szasz is making the policy argument that people without bodily lesions should, for that reason, not be subject to involuntary commitment, his illness analysis overcomplicates and obscures his point. What is required instead is an analysis of the connections (which are not obvious, to be sure) between lesions and the justifications and goals of commitment. Interposing the concept of illness between the concepts of lesions and involuntary commitment accomplishes nothing from an analytical standpoint and, predictably enough, has given rise to an utterly irrelevant and unproductive dispute between Szasz and his critics concerning the definition of illness and the function of lesions in that definition.<sup>13</sup> Again, if we wish to determine whether the mentally disordered ought to be subject to involuntary commitment, we ought to ask whether any characteristics of mental disorders reasonably justify such action. We ought not to waste our time wondering what characteristics are possessed by illnesses and whether mental disorders share them.

A final example of the unnecessary or erroneous use of the illness or disease

---

<sup>12</sup>For a statement of Szasz's position on the physical lesion requirement and the metaphorical nature of mental illness, see T. SZASZ, *supra* note 5. For an example of Szasz's application of that position to involuntary commitment, see T. SZASZ, *THE THEOLOGY OF MEDICINE* 137 (2nd ed. 1988). Perhaps, rather than making a policy argument, Szasz is engaging in an exercise in statutory interpretation of the following sort: (1) Involuntary commitment statutes by their terms apply to certain people with mental illnesses (e.g., "dangerous" mentally ill persons); (2) All illnesses, including mental illnesses, involve lesions; (3) Whatever mental problems people may have, none of them involves lesions; (4) No one has a mental illness (stated differently, mental illnesses do not exist); (5) The involuntary commitment statutes apply to no one and should be repealed. If this is the case, Szasz is ignoring a fundamental rule of statutory interpretation, namely, that a statute should be construed so as to effectuate the intent of the legislature enacting it. Legislatures passing these statutes undoubtedly had a more or less clear idea of a class of people to which they were to apply, and it seems quite unlikely that any such legislature intended to impose a "lesion" requirement for applicability of its statute.

<sup>13</sup>E.g., R. ISAAC AND V. ARMAT, *MADNESS IN THE STREETS* 33-41 (1990); T. SZASZ, *supra* note 5, at 156-58; Moore, *Some Myths About "Mental Illness"*, 32 *ARCHIVES GEN. PSYCHIATRY* 1483 (1975); Pies, *On Myths and Countermyths*, 36 *ARCHIVES GEN. PSYCHIATRY* 140 (1979); Schoenfeld, *An Analysis of the Views of Thomas S. Szasz*, 4 *J. PSYCHIATRY AND LAW* 245 (1976). Interestingly enough, even as sophisticated a critic as Moore, who in another context expressly acknowledges that "no definition of mental disorder for the psychiatric profession can or should be controlling of other disciplines' definitions of the phrase" (Moore, *supra* note 5, at 88) and that each discipline "must govern its definitions by its own purposes," (*id.*) finds it necessary to dispute Szasz's definition of "illness" and to offer a general one of his own ("being ill seems to involve something like being in a state of pain or discomfort, which, if not removed, may lead to premature death, and which, for its duration, incapacitates the patient from certain activities thought normal in our society." (Moore, *supra*, 32 *ARCHIVES GEN. PSYCHIATRY*, at 1490)) whose applicability to many mental disorders is questionable.

concept can be found in Nancy Andreasen's *The Broken Brain*.<sup>14</sup> There, Andreasen first concludes that "[t]he major psychiatric illnesses are diseases" which "are caused principally by biological factors. . . ." <sup>15</sup> She then states what she calls an implication of this biological model: "Mental illnesses are not due to 'bad habits' or weakness of will. *Because they are illnesses, they cannot be cured by acts of will*" [italics added].<sup>16</sup> At another point, she states:

Psychiatry now recognizes that the serious mental illnesses are *diseases* in the same sense that cancer or high blood pressure are diseases. Mental illnesses are diseases that affect the brain, which is an organ of the body just as the heart or stomach is. People who suffer from mental illness suffer from a *sick or broken brain*, not from weak will, laziness, bad character, or bad upbringing.<sup>17</sup>

Andreasen thus appears to determine that mental disorders are diseases because they are biologically caused. And, because mental disorders are diseases, she argues, they have no intentional or moral dimension. But if biological causation implies unintentionality or amorality (and such an implication is by no means obvious or unproblematical, see discussion *infra*), mental disorders have those qualities because they are biological, not because they are diseases. By arguing from biology to disease to unintentionality, instead of directly from biology to unintentionality, Andreasen, completely unnecessarily, raises the complex question of the proper definition of *disease* and exposes her argument to the obvious objection that biological causation cannot be the sole definitional criterion of disease.

Alternatively, if unintentionality and amorality depend on factors other than biological causation, Andreasen's argument from disease to unintentionality and amorality, where disease is defined only by biological causation, is fallacious.

If mental disorders are illnesses, it is because they possess all of the necessary defining characteristics of illnesses; thus to legitimately call mental disorders illnesses, we must know in advance that they possess those characteristics. This being the case, if our decisions about the mentally disordered are based on one or more of those characteristics (or their consequences), we can make those decisions before, and without, making the illness definition. And, to the extent that our decisions are based on characteristics unrelated to those that define *illness*, determining that mental disorders are illnesses is irrelevant to those decisions, for we cannot legitimately infer that a mental disorder, solely by virtue of its being an illness, possesses any characteristic other than those that define *illness* or that follow from such defining characteristics.

### **Why is the Illness Analysis Used?**

If our analysis up to this point is correct, why are so much time and effort devoted to the enterprise of determining whether mental disorders are illnesses?

<sup>14</sup>N. ANDREASEN, *supra* note 5.

<sup>15</sup>*Id.* at 29–30.

<sup>16</sup>*Id.* at 31.

<sup>17</sup>*Id.* at 8.



While there are probably as many explanations as there are participants in the debate, one can speculate that there are a few common reasons.

First, people may simply and mistakenly believe that the illness analysis is a valuable way of answering questions about the moral and legal responsibility of the mentally disordered, their proper treatment by society, and the appropriate type or types of health professionals to have jurisdiction, so to speak, over the mentally disordered. The illness analysis has been a very common way of approaching these questions and we may note that old habits of thought, even needlessly complex and confusing or fallacious ones, die hard. Further, despite its logical problems, the illness analysis seems to be a powerful rhetorical device. Many people, for example, appear unquestioningly to accept the notion that if a person's conduct was caused by an illness, he or she is not morally responsible, and should not be legally responsible, for that conduct, *even if* the bases upon which the person was found to be ill have nothing at all to do with questions of responsibility.<sup>18</sup> Why this should be so is not clear to us and, in any case, is beyond the scope of this paper.

A second possible reason for the persistence of the illness debate is that its outcome may have far-reaching economic effects. The status of mental disorders as illnesses, diseases, sicknesses, or disabilities could, for example, result in their being covered under health insurance policies or triggering certain statutory disability or welfare payments.<sup>19</sup> The economics of, and intent underlying, insurance policies and disability and welfare laws are beyond the scope of this paper. In passing, however, we would like to assert that it is just those sorts of factors (i.e., economics and the intentions of the parties to the insurance contracts and of the legislatures in enacting such laws), rather than whether mental disorders come within some textbook definition of *illness* unrelated to those factors, that should determine such issues.

Third, it might be argued that, regardless of its analytic value (and, perhaps, regardless of its truth value), the characterization of mental disorders as illnesses serves important practical purposes, namely, it encourages mentally disordered people to seek treatment and frees them from feelings of guilt that they are themselves to blame for their difficulties. That the fulfillment of these purposes (if, indeed, using *illness* language does so) is a good thing for society, and not merely for the mental health profession, is by no means clear. For every troubled person who truly benefits from obtaining mental health care,

---

<sup>18</sup>*E.g., id.*

<sup>19</sup>*See, e.g.,* Kitchen v. Time Ins. Co., 232 N.W. 2d 863 (Iowa 1975), where the court determined that alcoholism was a "sickness," as that term was used in a health insurance policy, after quoting Driver v. Hinnant, 356 F.2d 761, 764 (4th Cir. 1966): "This addiction—chronic alcoholism—is now almost universally accepted medically as a disease"; McGarrah v. State Accident Ins. Fund Corp., 296 Or. 145, 675 P. 2d 159 (1983), where the court, in deciding that the claimant's stress-related mental disorder was a compensable "occupational disease" under a state workers' compensation statute, stated:

If the legislature wants employers and compensation carriers to be relieved from the burden of such claims and wishes to change the occupational disease law to exclude mental disorders, such as exhaustively set forth in the American Psychiatric Association's Diagnostic and Statistical Manual of Mental Disorders, then the legislature can amend the statute to exclude specifically compensation for mental or physical disorders arising from job stress and conditions.

*Id.*, 675 P.2d at 170 (citation and footnote omitted).

one can certainly imagine another who, for example, is led needlessly to seek treatment, or who pays too much in money or time to obtain it, or whose family or friends are made less happy by his or her therapeutic transformation. Furthermore, while freeing people from their feelings of responsibility for their actions may make them feel better about themselves (though it may not<sup>20</sup>), and may in some cases be morally imperative, it also serves to weaken perhaps the strongest social control on harmful or otherwise unacceptable conduct. Finally, a strong argument could be made that people ought to base their decisions about whether they are sick, need treatment and are responsible for their actions on rational analysis rather than on mental health profession public relations. In any case, this justification for the illness characterization expressly treats the analytic aspect of the mental-disorder-as-illness debate as irrelevant, and it is with precisely that aspect that we are concerned in this paper.

Finally, the question of whether mental disorders are illnesses may be a proxy of sorts for the more fundamental question of whether mental disorders are biologically caused. Again, the best way to answer that latter question would certainly appear to be to address it and the purported consequences of its answer directly, rather than to muddy the waters with the illness concept and its attendant confusions. In the next section of this paper, we discuss the concept and consequences of biological causation of mental disorders.

### The Organicity Controversy

We have shown that the argument over whether mental disorders are illnesses is misguided, and have recommended that, alternatively, debate should focus on the properties of mental disorders. One property which heretofore has been thought to be of decisive importance by both opponents and advocates of the illness position is the degree to which mental disorders are biologically caused. Szasz, for instance, denies that there is reliable scientific evidence for an organic basis to mental illness.<sup>21</sup> The denial of an organic basis is central to his claim that mental illness is a "myth," which in turn is one of the bases of his objections to current psychiatric practices. Szasz's critics counter that he ignores overwhelming evidence that at least some mental disorders, such as schizophrenia, have an organic basis.<sup>22</sup> Both sides appear to believe that the degree of evidence for biological causation of a mental disorder is an important consideration in the question of how those with the mental disorder should be treated by society. This consensus is mistaken. We contend that as with the illness debate, the biology debate is misplaced. The question of whether mentally disordered behavior is biologically caused has no obvious implications for the resolution of the aforementioned controversial social and legal issues. The

---

<sup>20</sup>See Warner, Taylor, Powers and Hyman, *Acceptance of the Mental Illness Label by Psychotic Patients: Effects on Functioning*, 59 AM. J. ORTHOPSYCHIATRY 398 (1989).

<sup>21</sup>Szasz asserts the necessity of discovering a "lesion" in the brain which correlates highly with a behavioral syndrome. See T. SZASZ, *supra* note 5. This would be sufficient but is not necessary to establish a biological difference between members of two classes (e.g., mentally disordered versus mentally normal). There are other ways of establishing such a difference which require no laboratory evidence whatsoever. For instance, certain patterns of familiarity suggest a genetic basis for a condition.

<sup>22</sup>E.g., R. ISAAC AND V. ARMAT, *supra* note 13; Schoenfeld, *supra* note 13.

belief that biological causation has such implications is based primarily on mistaken notions of determinism and free will.

### **Current Status of the Evidence for a Biological Role in Mental Disorders**

Before we attack the arguments which have been made from the premise that mental disorders have an organic basis, we briefly consider the necessity of doing so. For if Szasz is correct that there is no convincing evidence for a physiological or organic basis of any so-called mental illness, then our main purpose here is merely an arid intellectual exercise.

A behavioral trait may be said to have a biological, organic or physiological basis if a neurophysiological explanation exists for the trait supported by findings that on average, there are organic or physiological differences between the class of those who have the trait and the class of those who do not. For mental disorders, two kinds of evidence of organicity have been offered: differences in brain physiology<sup>23</sup> and genetic studies. Szasz focuses almost exclusively on the former, insisting that no brain lesion has been found that is necessary and sufficient for the pathogenesis of any so-called mental illness.<sup>24</sup> Although this is true, Szasz's criterion is unnecessarily stringent. Average differences between schizophrenics and normals have been demonstrated for a variety of phenomena, such as enlargement of the brain's ventricles and the number of dopamine receptors in certain brain areas.<sup>25</sup> The genetic evidence is still more impressive. Twin studies and adoption studies have repeatedly shown a strong genetic basis for individual differences in schizophrenia and manic depression.<sup>26</sup> It is true that no satisfactory biological explanation or diagnostic test as yet exists for schizophrenia or any other mental disorder. But there is every reason to believe that more adequate explanations and tests will be available in the future for at least some psychodiagnostic categories. Thus, to the extent that Szasz's arguments rely on our present ignorance, they are likely to become increasingly vulnerable.

### **Causation and Human Behavior**

Before we show that biological explanations have no implications for either policy or morality, we assert two premises on which our arguments depend:

1. Every event and every state of affairs (including every human action and every human thought, emotion and other mental state) is caused by, and is in fact a necessary consequence of, other prior events and states of

---

<sup>23</sup>Differences in brain physiology include phenomena as diverse as neurotransmitter differences and differences in brain anatomy. See I. GOTTESMAN, *SCHIZOPHRENIA GENESIS* (1990).

<sup>24</sup>See, e.g., T. SZASZ, *INSANITY: THE IDEA AND ITS CONSEQUENCES* (1987).

<sup>25</sup>See, e.g., Seidman, *Schizophrenia and Brain Dysfunction: An Integration of Recent Neurodiagnostic Findings*, 94 *PSYCHOLOGICAL BULLETIN* 195 (1983).

<sup>26</sup>See *id.* Although some scientists assert that the genetic evidence is fatally flawed (e.g., R. LEWONTIN, S. ROSE & L. KAMIN, *NOT IN OUR GENES* 197-231 (1984)), they are increasingly seen as unreasonably critical and politically motivated.

affairs, which themselves are caused by and are necessary consequences of prior events and states of affairs, and so on. In the realm of human behavior, this premise entails the denial of the existence of free will in its most fundamental sense. That sense of free will requires that the individual have a meaningful choice among alternative actions or decisions, each one possible. This is inconsistent with the notion that every event and mental state (including every apparent decision or choice) is completely determined by prior events and states of affairs.

2. All events and states of affairs occur in the physical world only. Thus, every cause and every effect is a physical occurrence or state. More specifically, all behavior is (most proximately) caused by neurophysiological states, usually in the brain, and those neurophysiological states are in turn caused by other physical events or states.

These two premises are versions, respectively, of the philosophical positions of determinism and materialism. They are the subjects of centuries-old controversies among philosophers, each having generated an enormous literature. Though our subsequent arguments rest heavily on the two premises, we are content merely to assert their validity. Both have been ably defended by more accomplished philosophers.<sup>27</sup> Our confidence in materialistic determinism stems in part from the near-miraculous success of the natural sciences, including the neurosciences. The past century has seen a proliferation of scientific disciplines which have identified biological causes of human behavior; for example, behavioral genetics, which has shown that genetic differences contribute substantially to human behavioral differences; neurology and neuropsychology, which have localized many abilities, emotions, and behaviors to specific parts of the brain; psychoneuroendocrinology, which has illuminated the degree to which the development of behavioral tendencies, as well as the regulation of our day-to-day behaviors are hormonally influenced; and evolutionary biology, which has given plausible evolutionary accounts of much of human nature (including moral judgments).<sup>28</sup> In addition to the empirical support, our confidence in materialistic determinism is due to the absence of a plausible alternative. Those who reject our premises will not find the rest of our arguments persuasive. (However, they should be required to give a plausible account of free will or of behavior which is not neurophysiologically caused.) We are more concerned that those who accept our premises recognize the consequences of doing so, namely, that the mere discovery that a given behavior has a biological cause has no consequences for how that behavior should be viewed or treated, from a social or legal standpoint.

### **The Irrelevance of Biological Causation**

The irrelevance of the mere fact of biological causation is easily demonstrated by the following argument: By assumption, all behavior is neurophysio-

<sup>27</sup>For introductions (including bibliographies) to the philosophical disputes concerning determinism and materialism, see Taylor, *Determinism*, in 2 THE ENCYCLOPEDIA OF PHILOSOPHY 359 (1967), and Campbell, *Materialism*, in 5 THE ENCYCLOPEDIA OF PHILOSOPHY 179 (1967).

<sup>28</sup>For a survey of biological influences on behavior see N. CARLSON, *PHYSIOLOGY OF BEHAVIOR* (1991).

logically determined. To give a neurophysiological explanation of behavior is to give a biological account. Therefore, all behavior is biologically determined. If all behaviors share a given characteristic, that characteristic cannot provide a basis for differentiating among any behaviors in any way. In other words, if biological causation has implications for any behavior, it has identical implications for all behaviors, since all behaviors are biologically caused. Thus, the mere demonstration of a biological cause for a given behavior cannot mean that that behavior should be viewed or treated any differently than a second behavior for which no biological cause has been found since, by assumption, there are biological causes (though presently unknown) for the second behavior as well.

The above argument shows that biological causation in itself cannot be used to distinguish among behaviors in any way. The particular distinction often drawn, contrary to our argument, is based on the view that if an act is shown to be biologically caused, it is, for that reason, not the product of the actor's free will. These acts are, presumably, to be contrasted with acts for which no biological cause has been shown and which therefore might be the products of free will. A common philosophical position is that if an act is not the product of free will, then the actor is not responsible for it, and should not be held so.<sup>29</sup> According to this position, if a man had no choice in committing a biologically compelled insane act, then it is unfair to blame or punish him for it. If, through treatment, a mentally disordered woman can regain control over her own behavior instead of having it dictated biologically by her mental disorder, perhaps we are not depriving her of her liberty by committing and forcibly treating her, but are instead restoring her liberty.<sup>30</sup>

If people's wills were sometimes free, then it would be quite relevant whether a cause could be discovered for a particular act. If a cause were demonstrable, this would mean that the act in question was not free. (Failure to demonstrate a cause would not necessarily have any implications, since one might be investigating the wrong cause.) However, by our first premise, free acts—acts which are caused by the free will of the actor—do not occur. Therefore, it is not necessary to discover the cause of an act in order to conclude that the act was not free. However interesting or important on other grounds the specific cause of the act might be (and, of course, such knowledge could be of crucial importance in understanding and treating mental problems), this knowledge is superfluous to the question of the existence of free will, which we answer negatively on a priori grounds.

The argument of this section is quite simple: Distinctions among behaviors cannot be drawn on the basis of whether such behaviors are biologically caused, since all behaviors are biologically caused. In particular, any distinc-

---

<sup>29</sup>For a discussion of this and related positions see Taylor, *supra* note 27.

<sup>30</sup>Although we are concerned here with those who distinguish biologically caused behavior from free behavior, the biological level of causation deserves no precedence over other, less proximate, kinds of behavioral causes (e.g., social) in the diminution of freedom and responsibility. Morris, for instance has argued that poor socioeconomic status is just as good (or bad) an excuse for criminal acts as biological factors. N. MORRIS, *MADNESS AND THE CRIMINAL LAW* 62 (1982). If a behavior which is caused is not free, and if unfree acts should have different consequences from free acts, then any cause which might be demonstrated (though all causes are ultimately neurophysiologically mediated) is equally relevant.

tion between biologically caused, and therefore unfree, behaviors and behaviors which are not biologically caused, and therefore might be free, is mistaken because no behaviors of the latter variety occur. Therefore, the presence or absence of a biological explanation for a behavior cannot be used (a) as a basis on which to infer any properties about that behavior which distinguish it in any way from any other behavior, or (b) as a basis upon which to attribute to that behavior a property which no behavior possesses, that is, that the behavior is caused by free will.

At this point, we would like to address two possible objections to this argument. First, we shall address the question of the apparent inconsistency of determinism and moral responsibility. Second, we shall discuss the argument that while behaviors may indeed be indistinguishable on the basis of the *fact* of biological causation (since all behaviors are biologically caused), relevant social or legal distinctions may plausibly be made on the basis of different *types* of biological causation.

### **Determinism and Moral Responsibility**

An objection that may be raised to determinism is based on the notion that, as discussed above, determinism is inconsistent with free will (in its fundamental sense of the individual's having a meaningful choice among more than one possible alternative action or decision). If this is the case, the argument goes, the truth of determinism negates the possibility of moral responsibility, since we cannot attribute moral responsibility (whether it be praiseworthiness or blameworthiness) to an actor for a given act unless he or she was free not to do it. That is, it is said, we cannot rationally praise or blame people for actions that they had no choice but to do. Since determinism entails that every actor has no choice but to do everything he or she does, determinism makes moral responsibility for any act impossible. Our strong individual and societal commitments to moral responsibility may therefore lead us to reject determinism.

The problem of reconciling determinism and moral responsibility has vexed philosophers for centuries and, needless to say, we do not presume to offer a solution here. However, because determinism plays a central role in many of the arguments of this paper, we offer the following comments which we believe point toward some plausible grounds for questioning the apparent mutual exclusivity of determinism and responsibility, and the necessity of rejecting determinism if the two do indeed conflict.

1. It may very well be the case that, as an empirical matter, the moral distinctions and attributions of moral responsibility that people actually make can be explained by reference solely to the *type* of causal factors contributing to the act in question, and without any reference whatever to *whether or not* the act was fully determined by prior causes. That is, perhaps what is crucial to people's moral feelings about a given act is what caused the act rather than whether the actor was ultimately free not to do it. Consider two acts: (A) Ed robs a bank because he wants money in a hurry, and (B) Mary robs a bank because she is kidnapped by an

Animal Rights terrorist group and forced at gunpoint to do so. The behavior is the same in both cases. Both actions were determined. In the former case, the primary determinants include Ed's impulsive, antisocial disposition (originating in large part from his genetic inheritance and his criminal father's role modeling), his crack cocaine habit, and his landlord's insistence on payment of the late rent. In the latter case, the main cause is Mary's reluctance to become a martyr for laboratory rodents (which is largely a product of her evolved self-preservation instincts). We assume, of course, that all of these causal factors consist in states of Ed's and Mary's brains. Both actions are fully determined and—at least theoretically—completely explicable in terms of prior events and neurophysiological states. Neither act was committed freely in the sense of free will as undetermined behavior. Nevertheless, most people (including, presumably, Ed and Mary) would grant a profound moral difference between the two actions, with Ed being judged more severely. There is some generally accepted sense in which he is more responsible, and his actions freer (though, according to our determinism premise, never free to have acted differently). Let us refer to this sense, which appeals to our intuitions, as “weak free will and responsibility,” weak in that it does not imply that the actor was free in the strongest sense, i.e., that he could have chosen to act differently.

We do not intend to give an analysis of why the two examples differ with respect to some alternative weak sense of free will and responsibility. Nor do we offer any definition for those alternative senses. Our purpose here is to illustrate that despite determinism, moral distinctions assert themselves on our consciousness, suggesting that there may be some commonly understood sense of both free will and responsibility *which is independent of the degree of causal determination*. It is possible—indeed, we think it likely—that no one simple, easily articulated rule exists for explaining how people in fact classify acts and actors according to their degree of weak free will and responsibility. Rather, several interrelated factors doubtless contribute, possibly including the degree of duress on the actor; his deterrability, rationality, and determination to be malicious (e.g., as indicated by the degree of premeditation); mitigating factors such as provocation; and the opportunity the actor has had to adopt acceptable behaviors (e.g., the training he received from his parents).

Thus, one can both accept determinism and the negation of free will in its fundamental sense, and still distinguish among acts and actors on commonly accepted, widely shared moral grounds. And, in fact, many philosophers have concluded that determinism and moral responsibility are not incompatible, often by positing a sense of *free* that is consistent with determinism and comports at least roughly with common moral judgments. (Classic formulations of this position can be found in the work of Hobbes and Hume.<sup>31</sup>) This of course does not prove (nor is it

---

<sup>31</sup>For a discussion of the relevant positions of Hobbes and Hume (with a bibliography), see Taylor, *supra* note 27.

intended to prove) that fundamental free will is not a prerequisite for moral responsibility. We wish only to point out that ethical theories which do not presuppose fundamental free will and which nevertheless generate common moral judgments are possible and have been proposed by important philosophers.

2. The enormous utility of moral praise and blame in the social control of undesirable or unacceptable behavior is quite consistent with determinism and materialism—the actor's (completely determined) knowledge of the potential costs (blame) and benefits (praise) of his actions (which knowledge consists in a particular state of his brain) becomes a causal factor in his future conduct.
3. Of course none of this answers the question of whether it is possible, or meaningful, or fair to attribute moral qualities to behavior which is fully determined. Again, we do not have the answer to that question. It should, however, be kept in mind that if determinism and moral responsibility are indeed inconsistent (and many philosophers have concluded that they are not) so that one or the other must be rejected, it is by no means clear that determinism, and not responsibility, must go.

The notion that all of our actions are determined is a disturbing one which is inconsistent with our subjective experience as free agents. In contrast, the notion of moral responsibility is deeply rooted in our subjective experiences and intuitions. This is, perhaps, the reason that the inconsistency of the two doctrines (if they are indeed inconsistent) might at first blush appear to require the rejection of determinism. Clearly, though, our subjective experiences of ourselves as undetermined moral agents is of no logical relevance, and of no demonstrated empirical relevance, to the truth of either doctrine. Because we believe or feel something does not make it so, and, without an appropriate theory linking the fact of our belief to the truth of its content, does not even increase its probability of being so. We might add at this point that if the case for responsibility is to be based on our intuitions, distinguishing “responsible” acts from “non-responsible” acts on the basis of free will versus biological causation may well fare no better intuitively than determinism—biological causation has been implicated in many behaviors we intuitively may regard as paradigmatic moral vices, e.g., crime and drinking too much for one's own good.<sup>32</sup>

We believe that to the extent we experience our actions as the products of free will, we are simply mistaken. Furthermore, scientists are becoming increasingly successful at uncovering the causes of behaviors which most people would consider freely chosen. For example, identical twins reared apart show striking similarities in religiosity, traditionalism, and occupational interests.<sup>33</sup> And, as mentioned above, the progress of the natural

---

<sup>32</sup>Cloninger, *Neurogenetic Adaptive Mechanisms in Alcoholism*, 236 *SCIENCE* 410 (1987); Cloninger & Gottesman, *Genetic and Environmental Factors in Antisocial Behavior Disorders*, in *THE CAUSES OF CRIME: NEW BIOLOGICAL APPROACHES* 92 (1987).

<sup>33</sup>Bouchard, Lykken, McGue, Segal & Tellegen, *Sources of Human Psychological Differences: The Minnesota Study of Twins Reared Apart*, 250 *SCIENCE* 223 (1990).



sciences and the absence of a coherent alternative lend at least some degree of plausibility to determinism. The objective plausibility of morality, on the other hand, is highly questionable. Attempts to derive normative moral precepts from descriptive, factual, objective, or empirical premises have failed, and it must be said that there is no demonstrable objective basis for the validity of moral rules, and it does not appear that one is possible.<sup>34</sup> (This is not to say, by any means, that there is no objective *explanation*, e.g., evolutionary, for the existence of our moral beliefs; only that there is ultimately no objective ground for holding that one set of moral rules is, as a matter of normative ethics, correct and another incorrect.)

Thus, putting aside our intuitions, which are fundamentally irrelevant, it would appear that, if determinism and moral responsibility are indeed inconsistent, responsibility may very well be the better candidate, analytically, for rejection.

### Alternative Construals of Biological Causation

Distinctions in the social, legal, or moral treatment of various behaviors cannot be made according to the mere fact of causation, since all behavior is caused. However, it is possible that such distinctions can be made on the basis of the type of cause. Such distinctions cannot be made according to the mere fact of biological causation, since all behavior is biologically caused. There are, however, some uses of the term *biological* which do not apply equally to all behaviors, and hence, could conceivably be used to make social, legal, or moral distinctions. We distinguish three such alternative uses below. In each case, the designation *biological* is ill-advised. Nevertheless, each is commonly supposed to be synonymous with *biological*, and in each sense, at least some mental disorders have been shown to be "biological."

One distinction which might be thought to reflect the degree of biological causation is that some behaviors are caused by processes in the brain which are only ephemeral, whereas other behaviors are caused by more enduring processes (Distinction 1). That is, in the former case, there are only temporary relevant differences between the brains of those committing an act and those not doing so, while in the latter case, there are more permanent differences. For example, in the bank robbery case discussed above, some of the brain processes causing Ed to rob the bank are stable, such as the neurophysiological basis of his impulsivity. In Mary's case, the relevant physiological processes would disappear as soon as coercion stopped. Hence, if one examined brain scans of volunteer bank robbers and conscripted bank robbers, respectively, one would be more likely to detect abnormalities (i.e., differences

---

<sup>34</sup>As is the case with virtually all philosophical propositions, this position is subject to dispute. For an introduction (with bibliography), see Nielsen, *Problems of Ethics*, in 3 THE ENCYCLOPEDIA OF PHILOSOPHY 117 (1967), though in defense of our position we note that that article states that "philosophical attempts to set forth a rational and objective normative ethic have not been notably successful in spite of the fact that philosophers have been engaged in that activity for the last 24 centuries." *Id.* at 119.

compared to normals) in the volunteer bank robbers. Following conventional understanding, let us denote behaviors associated with more permanent brain states as more biological than behaviors associated with more transient states. (However, let us also keep in mind that this designation is inaccurate; transient brain states are in fact every bit as biological as more permanent ones.)

A related distinction is between behaviors which are learned versus those which are not (Distinction 2). This is different from the preceding distinction because it is generally accepted that learning produces long-term, essentially permanent, changes in the brain.<sup>35</sup> Thus, behaviors are to be distinguished according to whether they are correlated with brain processes which are due to learning. Since learned behaviors are commonly (though mistakenly) thought to be less biological than unlearned behaviors, we will retain that usage of *biological* for Distinction 2.

Finally, let us consider the well-specified distinction between genetically and environmentally determined traits (Distinction 3). There is a highly developed science aimed at measuring the relative contribution of genetic and environmental determinants of behavioral differences.<sup>36</sup> Conceptually, a behavioral trait is more genetic (heritable) if individual differences in the behavior remain high even if people share the same environment, whereas environment is more important if individual differences tend to disappear when people's environments are similar. The distinction between genetically and environmentally influenced behavior is of particular interest here since there is persuasive evidence that some mental disorders have a strong genetic component.<sup>37</sup> Following conventional (though, again, inaccurate) discourse, we refer to more heritable traits as being more biological than less heritable traits.<sup>38</sup>

The three distinctions outlined above are related, but not identical. For instance, behavioral traits which are highly heritable probably require some biochemical or structural brain mechanism which is inherited, thus linking Distinction 1 with Distinction 3. However, there can be genes for transient brain states as well.<sup>39</sup> Similarly, if two people behave differently because of different learning histories, then their behavioral difference is less likely to be a result of genetic factors, linking Distinctions 2 and 3. Finally, as noted above, as used in both Distinction 1 and Distinction 2 *biological* implies stable brain processes, but Distinction 2 biological behaviors are restricted to those stable brain processes which are not due to learning. If there is a common thread to the three distinctions, it is that each implies that a behavior is more biological if its likelihood depends more on qualities of the actor than on experience.

<sup>35</sup>Barinaga, *The Tide of Memory Turning*, 248 SCIENCE 1603 (1990).

<sup>36</sup>PLOMIN, DEFRIES & MCCLEARN, BEHAVIORAL GENETICS (1989).

<sup>37</sup>*Id.*

<sup>38</sup>Even on its own terms, the conventional understanding is problematic. On one hand, environmental events such as head injuries are biological, in a conventionally understood sense. On the other hand, genetic factors can operate through nonbiological routes. For instance, if physically unattractive people are more likely to be mistreated, and hence, to develop a mental disorder, this will primarily be attributed to genetic influence (since physical attractiveness is largely genetic).

<sup>39</sup>Identical twins show similar patterns of developmental *change* as well as continuity. Wilson, *The Louisville Twin Study: Developmental Synchronies in Behavior*, 54 CHILD DEVELOPMENT 298 (1983).

### Implications of Alternative Construals of Biological Causation

Do any of the three senses of “biological” elaborated above have any relevance to the determination of the social, legal, or moral consequences of an act? Since it is commonly believed that the more biological the cause, the less responsible the actor, we are especially concerned with the question of whether biological causation diminishes responsibility compared to nonbiological causation.

Clearly, if any explanation of an act is to be relevant to the degree of weak free will and responsibility involved, both of the following must exist: first, at least some partial theory about the nature of weak free will and responsibility; and second, evidence that the candidate explanation has implications for their presence or absence under the theory. For example, suppose that one could make a convincing argument that some minimal degree of rationality of the actor is a prerequisite for responsibility for an act. Suppose further that strong evidence existed that damage to the brain’s Area A made rationality impossible. One could then reasonably attempt to make the case that an individual who had a lesion in his Area A should not be held responsible for a bad act. Unfortunately, we are unaware of any attempts systematically to link a theory of weak free will and responsibility to a theory of the brain. Those who draw moral conclusions from biological causation write as if such conclusions are self-evident, and thus appear implicitly to be taking the position, which we have rejected above, that behaviors can be distinguished on the basis of the mere *fact* (as opposed to the *type*) of biological causation. Because we are neither moral philosophers nor neurophysiologists, we are not prepared to speculate about any hypothesis regarding the relationship between a particular type of brain process and moral responsibility, much less to offer a comprehensive theory. We can attest, however, that despite our best efforts we have failed to produce a satisfactory example of the relevance of any of the three senses of *biological* to any commonly accepted notion of weak free will.

Similarly, and more generally, if one wishes to argue that behavior caused by a particular one of these three types of biological factors ought to be treated differently, for *any* social or legal purpose, from behavior caused by other biological factors, one needs to have a theory justifying such differential treatment. For example, one theory might be that “genetic” conduct is not deterrable and therefore should not be punished under the criminal law. However, both premises of this theory ((a) genetic conduct is not deterrable, and (b) undeterrable conduct should not be punished) are questionable. In any case, it is not our purpose here to propose such a theory (we have none to propose) or to reject one. Our point is simply that if moral, social, or legal consequences are to be attributed to a particular *kind* of biological causation (as opposed to the general *fact* of biological causation), coherent distinctions among types of biological causation must be drawn and a plausible theory of the relevance of such distinctions to social, legal, or moral policy must be offered. Those who seek to draw policy conclusions from biological evidence typically do neither.

### Conclusion

Questions of how society and its laws ought to treat the mentally disordered are both important and vexing. We propose that such questions be addressed

head-on and answered by using the methodology appropriate to all issues of social policy—an explicit cost-benefit analysis of the alternatives, with costs and benefits conceived of broadly enough to encompass more than traditional economic factors, *e.g.*, moral considerations as well.

What is the proper role of the illness debate in this calculus? We believe it has none. The costs and benefits to the individual and to society, however broadly such costs and benefits are construed, are not affected in any way by the characterization of mental disorders as illnesses, or by the determination not to do so. At best, *illness* or *non-illness* is an ill-advised shorthand for a cluster of qualities which we must have known mental disorders possessed before we could legitimately decide that they were or were not illnesses. At worst, the analysis generates misleading and unfounded implications concerning issues of responsibility, involuntary commitment and areas of medical jurisdiction, and adds totally unnecessary layers of definitional complexity to questions that are certainly difficult enough in their most straightforward form.

The fate of the work of Thomas Szasz illustrates some of the harmful effects the illness analysis has had on the attempts of the medical, legal, and academic communities to answer policy questions concerning the mentally disordered. For approximately the past 30 years, in book after book, and article after article, Szasz has repeatedly (obsessively, some would say) presented very powerful ethical and policy arguments that so-called mental illness should not be considered an adequate justification to involuntarily confine or treat an individual or to absolve him from responsibility for his actions. These arguments are based on several factors, the primary ones being Szasz's contention (which, in our opinion, is irrefutable) that patterns of behavior determined to be mental illnesses are distinguishable from other patterns of behavior primarily on moral grounds (that is, patterns labelled *ill* are considered to be undesirable) and that in a nation as committed to individual freedom as the United States, society's disapproval of a person's behavior (so long as it is not criminal) cannot justify his incarceration and loss of other civil rights.

Unfortunately, Szasz's positions are based on another factor as well—his oft-cited and much-maligned argument that mental illness is a myth (*i.e.*, is not genuine illness). As discussed above, whether mental illnesses are actually illnesses is irrelevant to the social and ethical questions Szasz wishes to answer. Szasz's mental-illness-as-myth argument therefore cannot support his conclusions. What his use of the argument has done, however, is given his critics a weak point<sup>40</sup> at which to attack his ideas, though of course their attacks are no more relevant to his conclusions than the myth argument itself. As a consequence, Szasz's more probative and powerful arguments on these crucial issues receive far less attention than they merit and are often taken far less seriously than they deserve to be.

The illness method of analysis cannot be redeemed by focusing on the defini-

---

<sup>40</sup>Szasz's myth argument is a weak point for several reasons, among them his rather bizarre "lesion-as-illness" notion discussed above, and the fact that his "myth" terminology has given rise to the common misinterpretation (repeatedly denied by Szasz, *e.g.*, T. SZASZ, *supra* note 6, at 17; T. SZASZ, *supra* note 5, at 162) that he does not believe that the *conduct* we call mentally ill actually occurs.

tional quality of illnesses often thought to be of the greatest consequence, namely, biological causation. Since all behavior is biologically caused, attributing biological causation to a given type of behavior, by calling it an illness or otherwise, tells us exactly nothing about the social, moral, or legal implications that behavior ought to have. This leaves open the possibility that a particular *type* of biological cause for behavior (*e.g.*, genetic) may have legitimate implications of this sort. An adequate theory would, at a minimum, identify such a type of biological causation and demonstrate a legitimate social, legal, or moral consequence of that type of causation. We await the proposal of such a theory. In the meantime, it is abundantly clear that illness theories meet neither requirement.

We know that certain people behave in ways that might be characterized as abnormal, deviant, incomprehensible, irrational, dangerous, or disturbing. We know that this behavior, like all behavior, is biologically caused. What further knowledge do we gain by positing a definition of *illness* and determining that some or all of this behavior does or does not fit it? The answer is: none. Let us therefore leave behind the complexities of the illness analysis and the misguided attempts to distinguish among types of behavior on the basis of biological causation. Instead, using what we know about the characteristics of these types of behavior and the people who exhibit them, let us direct our efforts toward identifying, evaluating, and balancing the costs and benefits of various policies toward the mentally disordered so that we may treat them rationally, compassionately, and justly.