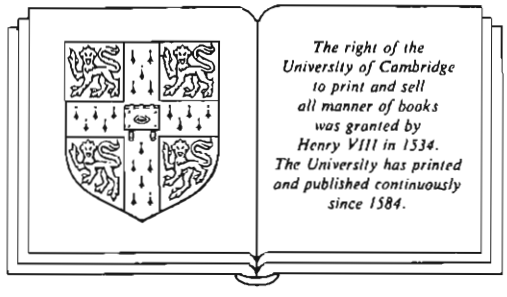
Law and Psychiatry

RETHINKING THE RELATIONSHIP

MICHAEL S. MOORE

University of Southern California Law Center



CAMBRIDGE UNIVERSITY PRESS

Cambridge

London New York New Rochelle Melbourne Sydney

The Insanity Tests and Their Underlying Moral Bases

A BRIEF SKETCH OF THE INSANITY TESTS

Insanity, as we noted in Chapter 2, was an excuse in most ancient systems of law. The two analogies that dominated the earliest ideas about insanity were the analogies of the insane to both children and to animals. In the nineteenth and twentieth centuries, however, four tests have dominated Anglo-American law on insanity. I shall briefly summarize these tests and then analyze what I take to be their underlying moral bases.

The M'Naghten *Test*

In 1843, Daniel M'Naghten shot and killed Edward Drummond, private secretary to the prime minister, Robert Peel. M'Naghten was under the delusion that he was being persecuted by a host of individuals throughout England and Scotland, including, he thought, the prime minister. At his trial he successfully raised the defense of insanity. Given the political unrest at the time, the suspicion that M'Naghten was merely feigning his illness, and the fact that she herself had recently been the target of an assassination attempt for which the assailant was excused by reason of insanity, the queen was outraged at the result of excusing M'Naghten from criminal responsibility. The House of Lords accordingly asked the judges to appear before them as a group and explain the proper tests of criminal insanity. The *M'Naghten* rules originated in the judges' answers to questions put to them by the House of Lords. The much-quoted and operative language of the judges' answers was that:

> to establish a defense on the grounds of insanity, it must be conclusively proved that, at the time of the committing of the act, the party accused was laboring under such a defect of reason, from the disease

> of the mind, as not to know the nature and quality of the act he was doing; or if he did know it, that he did not know what he was doing was wrong.[1]

The *M'Naghten* test quickly became the leading test for insanity in England and America. It is still the exculsive test for insanity in many American states and part of the test of insanity in many others.[2] The essential elements of the test, about which so much has been written, are three: first, that one suffer from a defect of reason; second, that such defect stem from a disease of the mind; and third, that one lack knowledge of some kind, either the knowledge of what one is doing or the moral knowledge that what one is doing is wrong.

The Irresistible Impulse Test

One of the persistent criticisms of the *M'Naghten* test in the nineteenth century, and continuing to this day, is that it does not excuse from criminal responsibility a large enough class of persons. More specifically, the *M'Naghten* test was thought not to excuse those mentally ill persons who knew what they were doing and that it was wrong, but who nonetheless, because of their mental illness, did not have the ability to control their behavior. The distinction often made was between cognitive incapacity and volitional incapacity. The irresistible impulse test was formulated as a response to this criticism of *M'Naghten*. Although no one case gives a classic definition of legal insanity as irresistible impulse, one of the leading cases defined the defense as follows:

> Did he know right from wrong, as applied to the particular act in question . . . if he did have such knowledge, he may nevertheless not be legally responsible, if the two following conditions concur: (1) if, by reason of the duress of such mental disease, he had so far lost the power to choose between the right and wrong, and to avoid doing the act in question, as that his free agency was at the time destroyed; (2) and if, at the same time, the alleged crime was so connected with such mental disease, in the relation of cause and effect, as to have been the product of it solely.[3]

As Abraham Goldstein has pointed out,[4] the common label for this test – "irresistible impulse" – is misleading. The essential notion of the test is that, because of mental illness, one has lost the power to control himself. The test might more properly be called a "loss of control" test.

The American Law Institute's Definition of Legal Insanity

In the 1950s the scholars and judges comprising the American Law Institute proposed a Model Penal Code for adoption by state and federal

jurisdictions. Section 4.01 of the code included a new definition of legal insanity: "A person is not responsible for criminal conduct if at the time of such conduct as the result of mental disease or defect he lacks substantial capacity either to appreciate the criminality of his conduct or to conform his conduct to the requirements of law."[5] This test, or some variant of it, has been adopted within the last twenty years by all of the federal courts of appeals and by a substantial number of states.

The New Hampshire and Durham *Experiments*

As I shall describe in much more detail shortly, the District of Columbia (between 1954 and 1972) and New Hampshire adopted as their criterion of legal insanity the following: "An accused is not criminally responsible if his unlawful act is a product of mental disease or of mental defect."[6] This test is traditionally analyzed as having two elements: first, that the accused be suffering from a mental disease or mental defect (mental disease being used in the test as a synonym of mental illness[7]); and second, that his criminal act be the product of that diseased or defective condition.

THE MORAL AND PSYCHIATRIC PARADIGMS UNDERLYING THE INSANITY TESTS

Writers in the area of legal insanity have sometimes conflated two quite distinct concepts: legal insanity and mental illness.[8] Although I shall argue ultimately that the two ought to be equated and legal insanity defined as mental illness (at least as the mental illness of the popular moral paradigm I analyzed in Chapter 5), it is nonetheless essential at this stage to keep the two distinct.

Some order can be brought into these tests by observing the relationship between mental illness and legal insanity. To be legally insane is to be excused from criminal responsibility. Each of these definitions of legal insanity thus is a test determining when an accused is or is not responsible in the criminal law. Each of the tests, it will be observed, has mental illness or some related concept as one of its elements. With the exception of *Durham* and the New Hampshire rule, however, the definitions of legal insanity typically do not equate insanity and mental illness, but use mental illness as only one element, one criterion, to determine when someone is legally insane. To be legally insane under *M'Naghten*, irresistible impulse, or the American Law Institute test requires, in addition, that certain other criteria be met. The imposition of these additional criteria should be seen as attempts by lawyers to relate psychiatric views about mental illness prevailing at the time the tests were adopted to well-established moral and legal paradigms of excuses from responsibility. To see this we must return to the taxonomy of excuses in Chapter 2.

In criminal law, as in morals, two general sorts of conditions excuse: ignorance that is not itself culpable, and compulsion. Such excuses are distinct from other modes of defeating the ascription of legal or moral responsibility known as justification, such as self-defense. These two moral excuses are as old as Aristotle and are embodied in contemporary criminal law.[9] If one makes a mistake of fact about a material element of a crime (e.g., one believes some substance to be harmless food coloring when, in fact, it is poison), one is not held liable to punishment; similarly, those who act under duress (threats by others), or those who act in response to their victims' provoking acts, are also fully or partially excused from punishment.

It is these two traditionally excusing conditions that have been adapted by lawyers and judges as they have added criteria to legal insanity beyond the requirement that an accused must be mentally ill. There are thus basically two kinds of traditional insanity tests: those based on the ignorance of the mentally ill accused person; and those based on some notion of his being compelled to act as he did.[10]

The *M'Naghten* formulation quite obviously is of the first type, which turns on the ignorance of the accused about what he is doing or its moral status. The *M'Naghten* opinion was written, and its rules adopted by other courts, at a time when delusions were thought to be the prominent symptoms of mental illness. Indeed, many judges and lawyers thought that the presence of delusions was the only criterion for being mentally ill. As Sir John Nicholl stated in *Dew v. Clark:*

> The true criterion, the true test of the absence or presence of insanity, I take to be the absence or presence of what, used in a certain sense of it, is comprisable in a single term, namely, "delusion". . . I look upon delusion in this sense of it, and insanity, to be almost, if not altogether, convertible terms.[11]

It is thus not surprising that a test should have evolved that combined this conception of mental illness—as delusion—with the long-existing moral paradigm that ignorance was an excuse. Although the old language about moral knowledge was retained, the rationale for the language was forgotten; as transformed by the *M'Naghten* rules, the knowledge required was not the general knowledge that was the measure of when a child or an insane person has the mental capacity to be treated as a responsible agent, but was rather the knowledge relevant to determining whether one could avail oneself of the excuse of ignorance of fact or law, an excuse available to sane as well as insane persons.[12]

The irresistible impulse test also represents this fitting of psychiatric insight into an already existing paradigm of moral excuse. It was principally Isaac Ray's criticism of the *M'Naghten* test in the nineteenth century, to the effect that delusion was not the only symptom of mental illness, that

persuaded courts to frame an alternative definition of legal insanity around the other existing paradigm of moral excuse, compulsion.[13] Whether a person had lost his ability to control himself, a "volitional incapacity" as opposed to a "cognitive impairment," was the essential question under the irresistible impulse test. Such a test received further impetus with the prevalence of psychoanalytic theories in the 1920s, whose talk of instinctual drives, energies, and forces all at least seemed to add to the idea that mentally ill persons are in some fundamental sense compelled to act as they do and are thus not responsible.

The American Law Institute's test is simply a rewording of each of these aspects of the nineteenth-century tests of legal insanity and their joinder into a single test. Instead of focusing on "knowledge," as in *M'Naghten*, the ALI test talks of "appreciating" the criminality of conduct; in place of irresistible impulses and inabilities to control, the ALI talks of a lack of substantial capacity to conform conduct to requirements of law. The moral paradigms invoked are wholly the same.

Each of these three tests shares a common and fundamental defect: They assume that legal insanity is an *excuse* for the particular acts done, not a general *status* attached to a class of human beings who are not accountable agents. Worse, they assume that insanity is not even a special excuse but is collapsible into the traditional excuses of ignorance or compulsion. There is, in such a view, nothing special about being crazy; one's responsibility is affected only if one can avail oneself of one of the two traditional excuses.

A similar misunderstanding pervades yet another test for insanity often proposed but not yet widely adopted. This view, often called the "elements" approach, also urges that mental illness is not an excuse. Rather, this view urges, it is only relevant to responsibility when it negates the *mens rea* element required for most crimes. Since the *mens rea* requirement for most crimes is intentionality (in Bentham's sense of purpose *or* knowledge), this view usually asserts that mental illness reduces responsibility only when it can be shown to have prevented an accused from forming the intention required for criminal liability. This view, although long accepted by some legal theorists and psychiatrists,[14] and although increasingly proposed as the best test for insanity, currently has its doctrinal home in the related diminished capacity defense.[15]

The problem with each of these views is that they fail to capture our moral intuitions about what it is about crazy people that precludes responsibility. Consider Daniel M'Naghten himself, who manifested some of the classic symptoms of paranoia and who would be considered in popular understanding as quite crazy. First of all, M'Naghten had the *intent* required for murder in England: He shot the gun with the purpose of killing another human being. True, he thought he was killing Prime Minister Peel

when in fact he was killing Peel's secretary, Drummond. But such mistakes about the identity of the intended victim never excuse in law, as the doctrine of "transferred intent" has long established. In every ordinary and legal sense of the word, M'Naghten *intended* the death of another. Similarly, he knew the "nature and quality of his act"; he knew its wrongfulness; he "appreciated its criminality." He made no *mistakes* about what he was doing—he knew he was shooting, and he knew that he was killing—nor was he ignorant of the legal and moral prohibitions against killing. Finally, there is no very persuasive case for saying that M'Naghten was *compelled* to do what he did. True, under the facts as he believed them to be, he had a hard choice to make. He believed that he was being persecuted by Peel and others and that if he did not strike first, he himself would be hurt or worse. Yet for a sane person such beliefs, even if true, would not give rise to any valid duress defense (there being no threats of immediate harm), nor can such preemptive strikes be justified as self-defense.

The short of it is that M'Naghten should flunk not only the test that bears his name, but all of the standard insanity tests with the exception of *Durham*. Yet I think our intuition is that someone like M'Naghten, who was very crazy, should not be responsible. It is true that we do get angry with even very crazy people when they do illegal acts. But we should not regard such anger as any more than our temporary emotional reaction to having been harmed, either actually or vicariously. After all, we also get angry with our children or our pets when they do things we dislike. Indeed, we even get angry at the chair on which we stub our toe. In all such cases, however, we should not mistake such anger for moral insight into such a being's responsibility. On reflection we should see that the insane, like the very young, are not sufficiently rational to be fairly blamed or punished. If this is so, then lawyers should give up their attempts to define legal insanity in a way that collapses it into some traditional excuse. Crazy people are not responsible because they are crazy, not because they always lack intentions, are ignorant, or are compelled.

The upshot of this is to incline one toward the New Hampshire and *Durham* formula of legal insanity, for this definition equates legal insanity with mental illness. Rather than incorporating long-existing paradigms of moral excuse from other areas of criminal law into a definition of insanity, the New Hampshire and *Durham* definitions regarded mental illness as itself an excusing condition, even if not accompanied by ignorance or compulsion.

Justice Doe, the originator of the New Hampshire test in the nineteenth century, thought that he was returning to the ancient ideas about legal insanity, namely, that mental illness itself excuses. In this respect he was correct; the status of being mentally ill, just like the status of being a child,

itself excuses one from responsibility. Unfortunately, however, Justice Doe of New Hampshire and Judge Bazelon of the District of Columbia Court of Appeals were both also heavily influenced by the psychiatric theories of mental illness with which they were contemporary. Their interpretations of their respective definitions of legal insanity were thus more influenced by the prevailing psychiatric theory than by the ancient paradigm to which, on occasion, Justice Doe thought he was returning.

The influence of psychiatry and contemporary opinions about mental illness on the judges who wrote the New Hampshire and *Durham* tests is complicated and will be treated separately below. The history of the origins and administration of the insanity test in those jurisdictions illuminates the influence psychiatric definitions have had on the legal definition of mental illness.

Two Experiments in Merging Legal and Psychiatric Definitions of Mental Illness

THE NEW HAMPSHIRE EXPERIMENT

In 1838, a little-known physician in Maine, Issac Ray, published *A Treatise on the Medical Jurisprudence of Insanity*.[16] The book quickly became the authority in its field and remained the seminal work in forensic psychiatry in the nineteenth century. The book was relied on heavily, for example, by the trial judge in M'Naghten's case (although ignored in the formulation of the *M'Naghten* rules in the House of Lords).[17] Ray's work was also highly influential on those early courts in America that adopted the irresistible impulse test.[18] More pertinent to our immediate concern, however, is Ray's direct and well-documented influence on the development of the definition of legal insanity in New Hampshire.

Ray, like most nineteenth-century psychiatrists, was convinced that mental illness was in essence a brain disease, and that lesions in the brain would eventually be discovered as the cause of mental disorders:

> It is undoubted truth that the manifestations of the intellect and those of the sentiments, propensities, and passions, or generally of the intellectual and affective powers are connected with and dependent upon the brain. It follows, then, as a corollary, that abnormal conditions of these powers are equally connected with abnormal conditions of the brain; but this is not merely a matter of inference. The dissections of many eminent observers . . . have placed it beyond a doubt; and no pathological fact is better established—though its correctness was for a long while doubted—than that deviations from the healthy structure are generally presented in the brains of insane subjects.[19]

This view of mental illness was eventually accepted by Justice Doe of the New Hampshire Supreme Court.

Early in his career Justice Doe sought out the views of Ray on insanity, via the good offices of a Dr. Tyler of Harvard Medical School. The three entered into a lengthy correspondence on the proper definition of legal insanity. Early in the correspondence, Doe asked Tyler:

> Is it now the settled opinion of the scientific world that insanity is only a physical disease, or the result of physical disease?[20]

If the answer were yes, Doe went on,

> Why should the court ever say to a jury more than this in cases of alleged mental aberration or active insanity: "If the disposition of property was the offspring or was caused or affected by mental disease, then it is not the will of the testator—the result of disease cannot have any effect in law."[21]

Both Doe and Ray believed mental illness to be a physical disease. From this premise, it was self-evident to Justice Doe that the following definition of legal insanity should be adopted (as it eventually was in New Hampshire):

> If the homicide was offspring or product of mental disease in the defendant he was not guilty by reason of insanity.[22]

The reasoning to this conclusion was based on a confusion continuing to this day: that if physical causes of behavior are discovered, the actor is *ipso facto* not responsible for that behavior. Doe thought that if the abnormal physical condition of a defendant's brain caused him to commit the criminal act, then necessarily his will, his power to choose, was extinguished; that is, it was not his act:

> For if the alleged act of a defendant was the act of his mental disease it was not in law his act, and he is no more responsible for it than he would be if it had been the act of his involuntary intoxication or of another person using the defendant's hand against his utmost resistance . . . [W]hen a disease is the propelling, uncontrollable power, the man is as innocent as the weapon—the mental and moral elements are as guiltless as the material. If his mental, moral and bodily strength is subjugated and pressed to an involuntary service, it is immaterial whether it is done by disease or by another man or a brute or any physical force of art or nature set in operation without any fault on his part.[23]

The confusion inherent in this chain of reasoning stems from the nature of the language we use to ascribe responsibility, in morals as well as law. As we discussed in Chapter 2, the words with which we describe human actions imply the causal agency of the actor; to be an action at all (as opposed to a mere bodily movement), we must be able to assert that the agent (or the self, the mind, the will, or what have you) was causally

responsible for the bodily movements. If we are told that something else caused the bodily movement, then it seems that the person was not causally responsible; for if some other set of conditions, such as an abnormal condition of the brain, was sufficient, that seems to imply that nothing else, such as the will of the actor, was even necessary.

The problem with such a view lies in the failure to see the possibility of there being differing sets of equally sufficient conditions existing to cause the same event. To say that a bodily movement is the product of an abnormal condition of the brain does not preclude one from describing that movement as an action performed by an agent for reasons. We have two vocabularies: that of movement and mechanical causation, and that of actions and reasons. Merely because scientists may discover lesions in the brain is not to preclude the application of the language of action and reasons. When there are mechanically caused movements there may nonetheless be intelligent actions.[24]

It is the language of actions and reasons with which the law deals. In ascribing responsibility, the law adopts with little change the conditions we all use to ascribe moral responsibility in everyday life, conditions framed in the language of action and intentionality. Once one perceives the "provisional independence" (discussed in Chapter 1) of this language from the language of natural science, then assumptions such as those of Doe and Ray become irrelevant to the proper definition of legal insanity. Mental illness may be a physical disease—that debate is certainly still alive—but for legal and moral purposes, the outcome is irrelevant. If mental illness excuses, it is not because it is the name of an as yet unknown physical cause.

Despite their now-questioned assumptions that mental illness is brain disease, and despite their further erroneous assumption that it is only because of such causal role that mental illness excuses, Doe and Ray were on the right track. For what they perceived to be inadequate about the *M'Naghten* test was its failure to recognize the responsibility-precluding nature of mental illness itself. Although for the wrong reason (physical causation), Doe and Ray believed that legal insanity should be equated with mental illness because mental illness by itself precludes responsibility. In that latter belief they were correct.

Their assumption that mental illness was a brain disease, however, led them into a further error for which they have been amply criticized by a century of legal scholars. Because of the assumed physical nature of mental illness, they thought that mental illness was a concept only scientists could meaningfully employ, and that accordingly, those scientists on the frontiers of investigating it—namely, psychiatrists—were the ones to inform the court on whether or not a particular defendant did or did not have those abnormal brain conditions amounting to mental illness. Again and again

define legal concepts in terms of legal purposes

in their correspondence and elsewhere, they reiterated the notion that the presence of mental disease was a fact to be testified to by medical experts. As Doe put it in an early dissent:

> Insanity . . . is the result of a certain pathological condition of the brain . . . and the tests and symptoms of this disease are no more matters of law than the tests or symptoms of any other disease in animal or vegetable life.[25]

In their correspondence, they argued:

> The law does not define disease—disease is so simple an expression that the law need go no further. What is a diseased condition of mind is to be settled by science and not by law—disease is wholly within the realm of natural law or the law of nature.[26]

Such views finally found authoritative expression in the New Hampshire Supreme Court's opinion in *State v. Pike*:

> The legal profession, in profound ignorance of mental disease, have assailed the superintendents of asylums, who knew all that was known on the subject, and to whom the world owes an incalculable debt, as visionary theorists and sentimental philosophers attempting to overturn settled principles of law; whereas in fact the legal profession were invading the province of medicine, and attempting to install old exploded medical theories in the place of facts established in the progress of scientific knowledge.[27]

The result of this view was clear: Psychiatrists would have the authoritative voice about who was legally insane and thus to be excused from criminal responsibility.[28]

The problem with such a result has been restated many times. The criteria of legal responsibility are for the law to settle. Even if mental illness is equated with legal insanity, its definition necessarily is a legal matter. It would be a pure coincidence if the concept of mental illness adopted by psychiatrists for the various purposes discussed in Chapter 5 were the same as the concept suitable to isolate the class of offenders who ought, consistently with the purposes of punishment, to be excused. Indeed, the point is quite general: The law must define legal concepts for itself in light of legal purposes. The law cannot simply adopt a concept developed by psychiatrists for therapeutic purposes, or for that matter any concept developed by any social scientists for explanatory purposes. The purposes of the law in question must govern the definition of any term appearing in that law; no other discipline's conceptualization can safely be adopted and plugged into a legal formula.

The psychiatrists at the state hospital in New Hampshire had the good sense to perceive that they could not have been asked under the New Hampshire test to incorporate their therapeutic notion of mental illness into the formula defining legal insanity. In fact, the concept of mental

illness they used in their testimony about criminal defendants was not the same as the concept they used in classifying their patients for treatment purposes. They had to take upon themselves the task Judge Doe and his successors on the New Hampshire bench should have undertaken; namely, to give a separate, legal definition of mental illness as a legally excusing condition.[29]

THE *DURHAM* EXPERIMENT

This fundamental lesson had to be relearned when the Court of Appeals for the District of Columbia decided *Durham v. United States* in 1954. The *Durham* rule regarding insanity, thought by the court to be "not unlike that followed by the New Hampshire court," was also that "an accused is not criminally responsible if his unlawful act was the product of mental disease or mental defect."[30] Unwittingly, the court followed in the steps of the New Hampshire court, both by relying initially on a medical conception of mental illness and by accepting the seeming consequence of the conception, namely, that causation was the real issue involved in the insanity defense. Each aspect of the court's reliance on the psychiatric paradigm of mental illness will be pursued separately.

Defining Mental Disease

Durham was decided explicitly to facilitate psychiatrists in placing their knowledge before the court, which they felt they could not do under the *M'Naghten* test. The influential Group for the Advancement of Psychiatry had earlier written a preliminary version of its report on criminal insanity, cited and relied upon in the *Durham* opinion. This report complained about "a barrier of communication which leaves the psychiatrist talking about 'mental illness' and the lawyer talking about 'right and wrong.' "[31] The test proposed by the committee, and in essence adopted in *Durham*, allowed psychiatrists to testify directly to the presence or absence of mental disease because the test was framed in terms of mental disease itself.

The problem, however, was the same as that which arose in New Hampshire almost a century earlier: Not every medically recognized mental disease could have been intended. Something more restrictive must have been intended by the phrase "mental disease or defect" in the *Durham* rule.

This problem became particularly glaring with regard to sociopaths, a diagnosis that had been applied to Monte Durham himself. Shortly after the *Durham* decision, the staff at St. Elizabeth's Hospital, which was composed of those psychiatrists most often called to testify in District of Columbia criminal cases, made a policy decision that sociopathic or psy-

chopathic personality disturbances would not be regarded as mental illnesses within the meaning of the *Durham* rule. Psychiatrists from St. Elizabeth's Hospital thereafter so testified in District of Columbia cases. Three years later, however, at a weekend meeting, the staff changed the policy, and decided that henceforth, psychopathic or sociopathic personality disturbances would be considered mental diseases for legal purposes. The Court of Appeals for the District of Columbia deferred to this psychiatric judgment, granting a new trial in one case involving a sociopathic individual because, having been tried before the change in classification by the psychiatrists, he was deprived of "new medical evidence . . . on an issue vital to his defense"[32] (namely, whether he was mentally ill). As Warren Burger, then a circuit judge who participated in that decision, later noted, "We tacitly conceded to St. Elizabeth's Hospital the power to alter drastically the scope of a rule of law by a weekend change of nomenclature."[33] This illegitimate transfer to psychiatrists of the power to decide the meaning of a legal rule on criminal responsibility resulted directly from the assumption of the District of Columbia judges that mental illness, as used in the rule, was the same concept as that used in medicine. This assumption, implicit in *Durham* itself, had been made explicit shortly after *Durham* was decided when the court of appeals held that

> mental "disease" means mental illness. Mental illnesses are of many sorts and have many characteristics. They, like physical illnesses, are the subject matter of medical science . . . . Many psychiatrists had come to understand there was a "legal insanity" different from any clinical mental illness. That of course was not true in a juridical sense. The law has no separate concept of a legally acceptable ailment which per se excuses the sufferer from criminal liability. The problems of the law in these cases are whether a person who has committed a specific criminal act . . . was suffering from a mental disease, that is, from a medically recognized illness of the mind.[34]

Perceiving that surely not every "medically recognized illness of the mind" excuses from criminal responsibility, psychiatrists in the District of Columbia, as those in New Hampshire before them, took it upon themselves to work out a legal concept of mental disease, first excluding, then including, sociopathic or psyschopathic personality disturbances.

Eight years after *Durham* was decided the court of appeals came to recognize that "what psychiatrists may consider a mental disease or defect for clinical purposes, where their concern is for treatment, may or may not be the same as mental disease or defect for the jury's purpose in determining criminal responsibility."[35] The court therefore attempted a legal definition of mental disease for the first time: "A mental disease or defect includes any abnormal condition of the mind which substantially affects mental or emotional processes and substantially impairs behavior control."[36]

The court of appeals thus finally undertook a task it should have undertaken originally: to give legal meaning to mental disease or defect as the phrase occured in the legal rule of responsibility. Unfortunately, however, the definition adopted is simply a regression to the more traditional types of definition of legal insanity. Mental disease is not itself actually defined, except insofar as a vague synonym is supplied: "abnormal condition of the mind." The informative part of the definition, which qualifies abnormal condition of the mind, is simply a reversion to those ancient moral paradigms already incorporated in the *M'Naghten* test, the irresistible impulse test, and the American Law Institute's test. Instead of "lack of knowledge" of the nature or quality of the act or its wrongfulness (*M'Naghten*) or the "substantial capacity to appreciate the criminality of conduct" (ALI), we have "substantially affects mental or emotional processes"; instead of language about acting under an irresistible impulse or lacking "substantial capacity to conform his conduct to the requirements of law" (ALI), we have "substantially impairs behavior controls." Instead of really giving a legal definition of mental illness, the court of appeals in *McDonald* abandoned the attempt by following the traditional formula for legal insanity: Mental illness, or some vague synonym for it, is only one element; added to it is some other traditionally excusing condition.[37] In doing so, the court of appeals abandoned the essentially correct insight behind the *Durham* case: that there is something about mental illness itself that precludes responsibility, irrespective of there being any ignorance about the nature of the particular crime or its prohibited nature, and irrespective of there being any excuse of compulsion.

The "Albatross" of Durham: *That Mental Illness Be the Cause of the Crime*

In 1972 the Court of Appeals for the District of Columbia abandoned the *Durham* rule entirely and adopted the American Law Institute's definition of legal insanity.[38] One of the principal reasons for doing so was the problem that court had been having with the "product" portion of the *Durham* rule: Psychiatric witnesses came to substitute their own judgments of the responsibility of the accused for that of the jury, and to phrase their conclusions on that ultimate issue in terms of whether or not the criminal act was the *product* of disease.

Contrary to some commentary on *Durham*,[39] this problem did not stem from the inherent lack of meaning that one can assign to the idea of mental illness causing crime. The concept of causation adopted by the court was unproblematic and consistent with the analysis of the causal relationship in

many areas of law and much of contemporary philosophy.[40] The concept of mental illness as naming the sort of thing that could properly be said to be a cause of behavior was more problematic; for Judge Bazelon, the author of *Durham*, clearly did not believe that mental illness was necessarily some brain condition that caused behavior in some mechanical way.[41] Presumably he and other members of the court meant to invoke psychological (as opposed to physical) theories of causation in thinking of the mental conditions of the mentally ill as a cause of their behavior. The court certainly invited expert testimony in terms of such psychological theories: "Description and explanation of the origin, development, and manifestations of the alleged disease are the chief functions of the expert witness. . . . The law wants from the medical experts . . . expert medical opinion as to the relationship, if any, between the disease and the act of which the prisoner is accused."[42]

The court got more than it wanted of such expert medical opinion. By 1967 it became clear to the court that psychiatric conclusions about "product" were often disguised moral judgments about the culpability of the accused, so the court flatly prohibited psychiatric testimony in terms of product, or cause and effect.[43] It never became clear to the court why psychiatrists were substituting their moral judgments for those of the jury, other than some speculations about defense counsel's strategies. For, if causation were a straightforward scientific question, why were the relevant scientists unable to testify to it without infecting their scientific judgments with their moral judgments?

The initial answer lay in the fact that causation was not an issue on which psychiatrists could be at all helpful, given the deterministic assumptions of modern psychiatry. For if psychiatrists, particularly those of a dynamic persuasion, took the causation test literally, then in every case in which a mentally ill person committed a crime they would have to find the crime to be caused by the disease. One accustomed to thinking in terms of the unified personality, of basic instinctual drives underlying all conscious motivations, of the pervasiveness of unconscious influence, and of the displaceability of psychic energy among seemingly unconnected objects could reach no other conclusion. Accordingly, any distinction psychiatric witnesses might make between mentally ill defendants would have to be made covertly on noncausal grounds.

What psychiatry essentially lacked—and still lacks—was any reconciliation of its own deterministic assumptions with the concept of responsibility. If mentally ill persons are excused because of their lack of "free will" (in a contracausal sense), then psychiatry could be of no help, for its theoretical commitment is that none of us enjoys the freedom the mentally ill are supposed to lack. Judge Bazelon himself ultimately came to wonder "how medical experts can be expected to provide information about the

impairment of free will, when free will would seem to be a philosophical and not a medical concept."[44]

The ultimate reason for the psychiatric unhelpfulness was that causal connection, in the sense used by the court, was not the issue anyway. Mental illness is not an excuse from responsibility because it causes criminal actions. This is true no matter if mental illness is construed as the name of some subset of physical causes themselves part of a general mechanistic account of human behavior – the nineteenth-century view – or if mental illness is thought to name some subset of psychological causes themselves part of a general, paramechanistic account of human behavior – a twentieth-century, Freudian view. In either case, the theories involved would excuse us all if such causation excuses the mentally ill. Psychiatric theories about causation, this time in their much less precise twentieth-century form of "psychic determinism," could accordingly only be a distraction from the actual criteria by virtue of which we separate the nonresponsible from the responsible in law and morals.

Had the court paid more attention to the meaning of mental illness and ignored any psychiatric theories of causation, it might have been able to have developed a meaningful idea of the relationship between the illness and the act. As is argued in Chapter 5, mental illness in ordinary understanding means an incapacity for rational action. One may do a certain act in a certain way because of such incapacity, but the "because" need not be construed on the model of mechanical or paramechanical causation.

A bridge may collapse because its materials lack the tensile strength to hold it up; a person may fail a test because he is stupid. In neither case have we cited an event, contiguous in space and time to another event, such that the first can be said to cause the second. The bridge's falling and the person's failing the test are events symptomatic of the general dispositional properties cited to explain them. Similarly, the relationship between illness and act the court might have sought was that the criminal act was symptomatic of the general incapacity for rational action of the mentally ill defendant. Because of its failure to work out an adequate legal definition of mental illness, however, the court could not reach this result but was bound to the psychiatric paradigm of mental illness it had adopted, that of a mechanical cause of behavior.

phrase could the wholesale adoption of the psychiatric definitions into law be appropriate. It is not impossible that there be such an overlap of legal and psychiatric purposes. If one's theory of punishment were purely rehabilitative, so that punishment was justified if and only if it was the best way of making an offender a better person, then there could be some overlap. For one of the purposes guiding psychiatric definition (to define conditions treatable by psychiatrists) could then also guide the definition of a class of offenders (the insane) who are not best rehabilitated by punishment. If punishment rarely cures, then anyone who can best be cured of his criminal tendencies by psychiatric treatment should, in a purely rehabilitative theory of punishment, be excused from punishment in order to obtain the more efficacious psychiatric treatment.

To see whether there really is any such overlap between the legal purposes in defining mental illness and the psychiatric purposes in defining the phrase, we must take a brief excursion into punishment theory. One can understand how mental illness should be defined in the criminal case law only if one first understands the purposes for punishing and for excusing from punishment.

Is One Responsible for Unconsciously Acting Intentionally or for Acting with Certain Unconscious Reasons?

RESPONSIBILITY AND CONSCIOUSNESS

In the preceding discussions the possibility has been left open that some of the phenomena classically cited as manifestations of the unconscious could

constitute unconscious actions, unconsciously intentional actions, or unconscious attempts. The examples given that might constitute such actions were (1) those dreams for which the extended memory can recapture the dreamer's supervenient attitude during the dream, that is, the belief that the dreamer was actively creating the dream; (2) those parapraxes, such as Freud's touching the girl, for which the actor's extended memory can recapture the belief necessary to consider the action as intentional; and (3) those neurotic symptoms, such as those of the girl of 19, that can perhaps be viewed as unconscious attempts because perhaps one can discover the belief necessary to complete the actor's practical syllogism. In each case, the main question was one of there being any evidence for the existence of the necessary beliefs.

Assuming for purposes of argument the necessary beliefs do at least sometimes exist, does this have as its consequence that one is responsible for such dreams, parapraxes, or symptoms, in the same manner that one is responsible for conscious actions, intentions, or attempts? Consider the example, discussed by Freud, of the Roman emperor who put to death a man who dreamed that he had assassinated the ruler.[80] Freud's intuition is here probably correct: The punishment was unfair because the man was not responsible for his thought. Yet why not? Assume that punishment is imposed only under those conditions in which fault is fairly ascribed. Specifically, only those thoughts that are actively called forth are punished, and thoughts that just occur to one are not. Assume further that this dreamer could recapture the memory that he had created the scenario in which the emperor was killed. With such assumptions, it might seem that the dreamer could fairly be punished for dreaming the prohibited thought because he has met those conditions under which fault is fairly ascribed.

The implication of responsibility is in fact not this clear. This becomes apparent on reconsideration of the ambiguity of the meaning of know discussed earlier. A person may have unconscious knowledge of a fact and still truthfully assert that he does *not* know that fact. Statements such as "He knew that fact, but did not know that he knew it" make perfectly good sense, as Freud himself recognized. Such expressions make good sense because the speaker has varied the sense of know, using it in differing senses in the same sentence. One sense is synonymous with being aware of, whereas the other is not.

Because the concepts of action, intention, and reasons depend upon the beliefs of the actor, the ambiguity of believe will infect these concepts as well. Showing that an "unintentional forgetting" is intentional, as Freud purports to do,[81] is not to show that it was also not unintentional. If an unconscious intention is discovered, then the action will be both intentional and unintentional. This is not a contradiction; it means that the actor both knew and did not know that he was performing some action.[82]

It is simply a mistake to regard unconscious actions or intentions, even when they are truly discovered, as contradicting commonsense descriptions of nonactions or unintentional actions. Psychoanalytic insight about unconscious intentional actions can supplement one's commonsense descriptions in terms of consciously unintentional accidents, but it cannot force one to withdraw the latter. A forgetting remains unintentional (no awareness) even if it could be shown that that forgetting was also intentional (a belief recapturable by extended memory).

The result is that one is in a moral quandary about responsibility when (if ever) there are the necessary unconscious beliefs. The moral principles under which one ascribes responsibility seem to require *both* that: (1) if a person's movement is an action, and it is intentional, then he is responsible; *and* (2) if it is not an action, or if it is unintentional, then he is not responsible. Yet these movements for which the required unconscious beliefs do exist are both properly described as intentional actions and either as nonactions or as unintentional actions.

To avoid an ambiguity in the term responsibility one needs some way of choosing one sense of knowledge as *the* sense relevant to responsibility assessments. Because one sense of knowledge means being aware of, such a choice squarely raises the question of whether consciousness is required for ascription of responsibility. Do only *conscious* actions, intentions, and reasons count in responsibility assessments? To answer this question requires some additional moral principle, for the principles heretofore relied upon, those requiring actions, intentions, and reasons, allow one to answer either way for this limited class of hypothesized examples.

The moral principle in question might be called a principle of consciousness. It would hold that in order to ascribe fairly responsibility to a person for causing a harm, he must have *consciously* acted intentionally, and to ascribe fairly responsibility to a person for attempting to cause a harm, he must have acted with that harm as his *conscious* reason. Using both ordinary senses of the word conscious, such a principle would have two parts: (1) The person must *have been* conscious in the sense of being awake; and (2) the person must have been conscious *of,* in the sense of being aware of, his acting, his intentions, or his reasons, if he is fairly to be held responsible for causing, or trying to cause, some harm.

One might further sharpen the second part of this principle by eschewing certain senses of conscious of. A principle requiring that the actor *think* of his actions, intentions, or reasons as he acts would be far too narrow. No such stream of consciousness or Joycean sense of conscious of can be at all plausible as a moral requirement, for it would exclude far too many intentional actions or attempts where one is perfectly content to ascribe responsibility. One does not, for example, consciously think or deliberate about going to lunch, yet doing so is an intentional action for

which one may justly be held responsible (in those hard-to-imagine circumstances where going to lunch turns out to be a bad thing to do). It is much more plausible to construe the proposed moral principle as using the dispositional sense of conscious of discussed in Chapter 3. In this sense one is conscious of going to lunch if one can state what one is doing when one's attention is turned to the subject. One need not be thinking about some fact to be conscious of it in this sense.

Having stated the moral principle, how does one establish it as correct? There are two ways of noninstrumentally justifying a moral principle: either by a series of particular thought experiments, all of which point to the principle as the best general expression of the intuited results of those particular experiments, or by resort to a more general moral principle from which the principle to be justified is implied. If one's moral sense is fully coherent, the results of the particular thought experiments will match the implications of one's most general moral principles.[83] If the principle to be justified is included in that cohered set of judgments and principles, it will be justified in the only way in which moral principles can or should be justified.

We have already reviewed a set of particular thought experiments with which we may test the validity of the principle of consciousness. In those comparatively few examples of dreams, parapraxes, or symptoms that could possibly be thought to constitute unconscious actions, intentions, and reasons, one need only imagine that the consequence brought about or attempted was morally bad. Freud's example of the traitorous dream will serve as such an experiment, as will his touching the girl. If one imagines the physical separation of the parents of the girl of 19 to be a bad thing, her unconscious attempt to bring it about will also serve to isolate intuitions about responsibility.

Surely one's intuitions in the first of these three cases are the same as Freud's,[84] namely, the dreamer does not deserve to be executed because he is not responsible for the unconsciously produced dream thought of the emperor's death. Of course, one's intuition is to some extent influenced by two extraneous factors: (1) One is reluctant to punish persons for *any* thoughts unaccompanied by physical action, no matter how much such thoughts constitute mental acts and not mere happenings; and (2) one's principles of proportionality are violated even in a moral system that sanctions punishment for mental acts because execution is too severe a penalty for merely thinking of another's death. One removes these extraneous considerations by asking whether the dreamer can fairly be punished at all for his dreams, assuming that one is generally willing to punish persons for mental, as well as for physical, acts. Even so limited, is not one's intuition the same as Freud's?

Similar results are reached in examples such as Freud's sexual advance

and the 19-year-old girl's attempt to separate her parents. Neither of them quite literally knew what they were doing. It is true that if one allows that they had the unconscious beliefs Freud claims for them, there is a sense in which they did know what they were doing. But is it not the morally persuasive rejoinder that they did not know that they knew, and that their ignorance should excuse them even in the face of such unconscious knowledge?

Further particular examples may be found in the law in which both parts of the principle of consciousness seem to be clearly established. Unconsciousness, in the sense of a lack of *being conscious,* is a complete defense to both civil and criminal liability. Thus, if one is asleep, dazed by a blow on the head, put out by the shock of being shot in the stomach, or unconscious because of drugs or alcohol involuntarily ingested, one is not held legally liable for the harm one has caused.[85] Sometimes this may be reflective of the fact that the behavior in question is not an action, not even an unconscious one. Yet even if the actor's extended memory could recapture beliefs necessary to make out such behaviors as actions for reasons, he is not legally liable to either tort or criminal sanctions.

Similarly, if a person is conscious (awake) but is not conscious (aware) of his actions, his intentions, or his attempts, he is not subject to punishment or civil damages for those crimes or torts requiring that we act intentionally or for certain reasons. Thus, a person who acts under posthypnotic suggestion, and is not conscious of his bodily movements as an act he is doing, is not held responsible for any harm that ensues, even if we were willing to say that these were unconscious actions.[86] A person who only subliminally perceives some fact, such as another's copyrighted music, and then intentionally but unconsciously copies that music, is not held liable for *willfully* copying another's copyrighted material because he was not conscious of his intention to do so.[87]

Such examples are presented, not because the law is self-justifying or incapable of incorporating moral mistakes, but because the law is a source of further instances in which one should ask whether the actor could fairly be held responsible. The answers that judges and lawyers have given are at least good evidence of *their* intuitions about the matter: To ascribe responsibility fairly is to require consciousness. If one's own intuitions are similar, one has good reason to subscribe to a similar principle.

The other noninstrumentalist way to justify a moral principle is by resort to more general principles from which the principle in question is implied. H. L. A. Hart once argued that "in most western morality 'ought' implies 'can,' and a person who could not help doing what he did is not morally guilty."[88] John Rawls, following Hart, regards this as *the* principle of responsibility from which more particular principles, such as those requiring actions, intentions, or reasons, may be derived.[89] Whether the lat-

ter claim is true is debatable, but the principle, in any case, is one most persons would accept as basic.

The truly debatable point, of course, comes in specifying that the principle *means* in more concrete detail. Using determinist assumptions about the universe, there is a sense of can in which no one can help doing what they do. This, however, cannot be the sense of can employed by Hart, else no one is responsible. Rather, the power or ability referred to is the power to give effect to one's choices. The principle of responsibility is part of a principle of liberty that places a high value on protecting people's choices from the unnecessary interference of moral or legal sanctions.[90]

Whatever else the principle of responsibility might include, it should include the power or ability to appraise the moral worth of one's proposed actions. A person has such ability only if he has moral and factual knowledge of what he is doing and is able to integrate the two to perceive the moral quality of his action.[91] A person who lacks this ability cannot fairly be blamed because, although he is acting intentionally, he does not know that what he is doing is wrong. This requirement is applied in the treatment of children and the insane as nonmoral agents because they have not had a fair opportunity to acquire such moral knowledge.[92]

When the factual knowledge is unconscious, which it is by hypothesis in the class of cases considered here, the ability to perceive the moral nature of one's actions is lacking. One cannot draw the moral conclusion because the factual knowledge necessary to make the inference is unconscious. Only when the factual knowledge is recaptured by one's extended memory does one have the ability to bring one's moral knowledge to bear on the particular action. That such unconscious beliefs can be recaptured only means that one could have, and perhaps should have, recaptured them. Not having done so, however, one did not possess the factual knowledge necessary for responsibility.[93]

There is a sense in which one could have acted otherwise in such cases, and a sense in which one could have drawn the necessary inferences from unconscious beliefs. However, this is the sense in which it is said that a person guilty of a negligent omission *could have* done other than he did. This does not mean that he had the same power to avoid doing evil as someone who consciously knows what he is doing. Rather, it means that he could have placed himself in a position to have had that power if he had paid more attention to himself and to the world around him. Insofar as we all have a duty to know ourselves – and morally if not legally we certainly have such a duty[94] – then we may in some cases be responsible for causing harm. The unconscious, in such a case, does not show us that we are responsible because we acted intentionally and for reasons; rather, the unconscious affects responsibility only in that its existence creates a duty to know it. Failure at such a duty may be a culpably negligent omission

that causes others, as well as ourselves, harm. Such responsibility for negligent omission is quite different from the responsibility for positive behavior, the subject of this chapter.

FREUD'S CONCLUSIONS ON RESPONSIBILITY

Do Freudians really want to claim more about an extended responsibility than has been discussed? Some contemporary Freudians plainly do,[95] but Freud's own position is not clear. His intuition about the responsibility of the dreaming traitor suggests he does not. There is, however, a lot of Freudian *talk* about responsibility that seems to point the other way.

For example, Freud concluded for several reasons that a person must assume responsibility for the content of his dreams.[96] First, there is a therapeutic concern: "Obviously one must hold oneself responsible for the evil impulses of one's dreams. What else is one to do with them?"[97] Second, Freud said one has no real choice about accepting responsibility:

> If I were to give way to my moral pride and tried to decree that for purposes of moral valuation I might disregard the evil in the id and need not make my ego responsible for it, what use would that be to me? Experience shows me that *I nevertheless do take that responsibility,* that I am *somehow compelled to do so.* The physician will leave it to the jurist to construct for social purposes a responsibility that is artificially limited to the metapsychological ego.[98]

Third, Freud elsewhere noted that it is "instructive to get to know the much trampled soil from which our virtues proudly spring."[99] Given one's search for such knowledge, "There seems to be no justification for people's reluctance in accepting responsibility for the immorality of their dreams."[100]

Freud's apparent position on one's responsibility for parapraxes is similar. Regarding forgotten appointments, he noted:

> There are two situations in life in which even the layman is aware that forgetting—as far as intentions are concerned—cannot in any way claim to be considered as an elementary phenomenon not further reducible, but entitles him to conclude that there are such things as unavowed motives. What I have in mind are love-relationships and military discipline. A lover who has failed to keep a *rendezvous* will find it useless to make excuses for himself by telling the lady that unfortunately he completely forgot about it. She will not fail to reply: "A year ago you wouldn't have forgotten. You evidently don't care for me any longer." Even if he should seize on the psychological explanation mentioned above . . . and try to excuse his forgetfulness by pleading pressure of business, the only outcome would be that the lady, who will have become as sharp-sighted as a doctor in psychoanalysis, would reply: "How curious that business distractions like

> these never turned up in the past!" The lady is not of course wanting to deny the possibility of forgetting; it is only that she believes, not without reason, that practically the same inference—of there being some reluctance present—can be drawn from unintentional forgetting as from conscious evasion.[101]

The keen-sighted woman, Freud seems to suggest, rightly holds her lover responsible for missing the appointment.

Last, consider Freud's discussion of Dora and her responsibility for her cough:

> I now return to the reproach of malingering which Dora brought against her father. . . . I was obliged to point out to the patient that her present ill-health was just as much actuated by motives and was just as tendentious as had been Frau K.'s illness. [Frau K. was Dora's father's mistress who faked illness in order to be with the father at a health resort.] There could be no doubt, I said, that she had an aim in view which she hoped to gain by her illness. That aim could be none other than to detach her father from Frau K. She had been unable to achieve this by prayers or arguments; perhaps she hoped to succeed by frightening her father . . . , or by awakening his pity . . . , or if all of this was in vain, at least she would be taking her revenge on him.[102]

Freud again seems to be saying that Dora's cough, although nominally not a basic action on her part, was nonetheless something she was doing; that through her cough she was also performing the more complex actions of frightening her father, awakening his pity, and taking her revenge; and that she is just as responsible for these actions as was Frau K., who consciously malingered in order to be with Dora's father. Again, for Freud, all of this seems to follow from the fact that Dora's coughing is to be explained by her unconscious mental states.

Despite all this, Freud was not really committed to such a radical extension of responsibility. His first argument regarding responsibility for dreams invites a confusion between the retrospective and the prospective senses of responsible that I distinguished in Chapter 2. In effect, Freud is saying a person must hold himself responsible for doing something about himself if his dreams reveal wishes that he does not like. This argument is applicable to all of psychoanalytic therapy: One must *accept* responsibility to make oneself better, and one *makes* oneself responsible for that task. Accepting such prospective responsibility, however, does not mean that one is retrospectively responsible for the production of an immoral dream, as one would be for a waking fantasy that one consciously called to mind.

Of course, in therapy the analyst may have to convince the patient that he is retrospectively responsible for some events in his past to get him to accept the task of improving himself in the future. Indeed, that seems to be precisely Freud's strategy with respect to Dora. In this,

however, he may be simply playing on the patient's feelings of guilt for her wishes. As noted in Chapter 2, there is a difference between *feeling* guilty about dreams and *being* guilty for them in the sense of being responsible for their production.[103]

Freud's second argument that we are responsible for our dreams invites this same confusion between being guilty and feeling guilty. Even if Freud is convincing when he says that we are *compelled* by our guilt feeings to accept retrospective responsibility, this does not mean that we *are* responsible. To determine whether we are responsible for dreams, we need to know something other than the psychology of guilt or the therapeutic strategies of analysts. We need to know if events such as dreams, forgotten appointments, or neurotic symptoms are truly intentional actions for which we can fairly be held responsible.

In his own therapy Freud himself on occasion worked against the feelings of responsibility of his patients. In his case study of the Rat-Man, for example, Freud reported his discussion with the Rat-Man regarding the latter's responsibility for the unconscious mental states that were a part of his character, as follows:

> In the further course of our conversation I pointed out to him that he ought logically to consider himself as in no way responsible for any of these traits in his character; for all of these reprehensible impulses originated from his infancy, and were only derivatives of his infantile character surviving in his unconscious; and he must know that moral responsibility could not be applied to children.[104]

On such occasions Freud himself separated the feelings of responsibility (guilt) experienced by his patients from their true responsibility.

Freud's third argument about responsibility for dreams invites a different confusion, this time regarding the *object* for which one is responsible. Freud's insight that one's virtues may spring from unconscious mental states of which one may not be overproud is an insight into one's character. The unconscious wishes expressed in a dream may well show one to be less moral in one's inclinations than one had thought. Insofar as one accepts the Aristotelian notion of being responsible for one's character, one is responsible for such wishes no less than for other aspects of one's character.[105] This is not to say that one is responsible for performing an immoral dream in the same way as one can be responsible for performing an immoral play; the responsibility Freud speaks of here is for character and not for the bad consequences of actions.

The case of the keen-sighted woman, of whose responsibility assessments about her forgetful lover Freud apparently approves, is somewhat different. There are a number of possibilities about what is being claimed by the disappointed woman who holds her lover responsible for having forgotten his appointment with her.[106]

First, she might be playing on her lover's guilt, just as a therapist might use his patient's guilt. She, like the therapist, would not really be claiming that her lover was responsible for an action designed to disappoint her. Her saying that he is responsible, however, triggers his guilt feelings for whatever (presumably nontherapeutic) end she has in mind. As previously stated regarding Freud's claims about one's responsibility for dreams, we must be careful to distinguish guilt (feeling responsible) from responsibility (being responsible).

Second, the woman, as Freud does with dreams, might be confusing prospective with retrospective responsibility; she may want her lover to accept responsibility for not meeting her on this occasion so that in the future he will make himself responsible for meeting her. His prospective responsibility for not forgetting is not to be confused with his having performed some past action for which he is retrospectively responsible.

A third possibility is indicated by Freud's suggestion that the military is as wise as the woman in its insights about responsibility for nominal accidents. Both military and civilian law often hold individuals responsible for phenomena like forgotten appointments because other negligent breaches of duty preceded the event in question. For example, one may be responsible for a car accident caused by one's epileptic seizure while driving; this is not because the epileptic movements are actions, but rather because an epileptic is negligent in driving at all.[107] Analogously, one might hold the lover responsible for not meeting the woman on the grounds that he did certain acts that made it unlikely that he would remember the appointment. For example, he may have gotten drunk. Notice, however, that such an account *does not* succeed because the person had "unavowed motives" transforming his unintentional forgetting into an intentional action that he performed; rather, it succeeds because it finds some earlier action of the actor that causes the harm for which he is held responsible.

The sense in which Freud probably intends to impute responsibility in this case is not adequately characterized by any of the three preceding analyses, for none of these ways in which responsibility might be attributed makes use of the most likely explanation for the lover's forgetting. This would be that the lover no longer loves the woman and, even though he may not realize this, the absence of emotion causes him to forget the appointment. To discover the kind of responsibility intended by Freud, it is important to see that the woman's frustration, disappointment, and anger at her lover do not arise from his supposed unavowed motives nor do they depend upon characterizing his forgetting as intentional. Rather, the woman is angry because her lover does not care for her anymore. His failure to arrive at the appointed time shows that he does not care, not that he adopted his behavior as a means to show the woman that he no longer cares. His emotion, or lack of it, explains his behavior, but does not

mean that he chose, even unconsciously, that specific behavior as the means of achieving some particular desire.

Consider the following analogous example: One raises one's voice *because* one is angry. Explaining one's behavior in this manner might be taken to mean that one raised one's voice to show anger, in which case the raising of the voice was an intentional action, done for reasons. Much more often, however, the proffered explanation is the following non-reason-giving account: One's anger caused one's voice to become louder, just as it may have caused one's heart to beat faster. One did not perform an action for reasons; these occurrences—the raised voice and the faster heartbeat—were simply caused by one's emotional state.

The nature of the lover's responsibility for missing the appointment can thus not be the responsibility we have for our intentional actions. One might say that the angry person and the lover are responsible for the behavior in question (shouting or not showing up) because they are responsible for their character, which includes certain emotions (or lack thereof in the case of the fickle lover). As stated earlier, everyone may be somewhat responsible for his own character. This includes the character of one's wishes and emotional states. In this attenuated sense one might then hold the lover responsible for the forgotten appointment, not because he did so intentionally or for reasons, but because his forgetting was an expression of a character for which he is responsible.

In neither law nor morals is this kind of responsibility for a harm to be equated with the responsibility that we have when that harm is caused by our intentional actions. Neither the expression of anger and desire nor the lack of emotion is an intentional action that we perform. They are events that happen to us, even though we may have had some choice in the past as to whether such events would be "in character." We may well shun the friendship or the company of those with bad character, but we should not blame them in the way that we blame those who intentionally cause harm or try to do so.

We should note that these distinctions dispose equally well of Freud's arguments to Dora that she was responsible for her symptoms and for trying to separate her father from his mistress. The responsibility need only be Dora's prospective responsibility to improve herself, based on her acceptance of responsibility as part of Freud's therapeutic strategy. Also, the therapeutic strategy is furthered if she feels responsible (guilty) for causing her father suffering and anxiety.

We may thus be said to be responsible for dreams, slips, and symptoms in a variety of ways, all of them different from the way in which we are responsible for intentional actions and attempts. Perhaps not all of Freud's conclusions about responsibility can be fairly reinterpreted in this way, but enough of them can be to raise serious questions about whether Freud

himself really believed that his "discovery of the unconscious" compelled the extended responsibility discussed throughout this chapter.

10 The Unconscious as the Source of a Decreased Responsibility

Do Unconscious Mental States Diminish Responsibility Simply Because They Are Causes of Behavior?

Dynamically oriented psychiatrists believe that unconscious mental state explanations excuse behavior because they think such unconscious mental states *cause* the bad behavior and that causation is an excuse. The testimony in *Pollard* suggests this in saying that Pollard was "governed by unconscious drives"[5] that exempt him from responsibility.

The problem with such a view of excuse is that of a reductio: If Pollard is to be excused simply because his behavior was caused by unconscious mental states, why are all actions not similarly excused? If all conscious mental life is determined by unconscious mental states, as many psychoanalysts believe, why is everyone not excused for all of his actions, seemingly the product of his conscious decisions but in fact determined by his unconscious mental states?

Some accept this conclusion, not as an absurd implication of a therefore unacceptable premise, but as an important insight about human beings. John Hospers, for example, believes psychoanalysis to have shown that one's "conscious will is only an instrument, a slave, in the hands of a deep unconscious motivation which determines his action."[6] In Hospers's view,

one is much like a "puppet whose motions are manipulated from behind by invisible wires, or better still, by springs inside."[7] The effect on responsibility is sweeping. "Criminal actions in general are not actions for which their agents are responsible; the agents are passive, not active—they are victims."[8]

The proper way to view this supposed vitiation of responsibility by the discovery of the unconscious is as part of a general philosophical position known as hard determinism. Determinism is the doctrine that every event (including every human action) has a cause; there are no "free" (in the sense of uncaused) events, in a deterministic world view. *Hard* determinism is the view that moral responsibility for an action cannot be justified if that action is caused. Hard determinism interprets Hart's principle of responsibility (mentioned in Chapter 9) to require contracausal freedom for persons before they can fairly be blamed for their actions, for how "could one have acted otherwise" if one's actions were fully determined?[9]

Psychiatrists, and dynamic psychiatrists in particular, are among the most vehement of hard determinists. Freudians especially believe that at last we have a theory bringing human behavior under the universal, causal laws distinctive of science: "For the first time in history," Karl Menninger assures us, "we have a logical and systematic theory of personality, an explanation of what human nature is and how behavior is determined and modified."[10] Given the prominence of unconscious forces in such a theory, free action (in the sense of uncaused action) is merely an illusion:

> Freud, in 1904, brilliantly demonstrated by analysis of slips of the tongue, forgetting and trains of association that what we call free will or voluntary choice is merely the conscious rationalization of a chain of unconsciously determined processes. Each act of will, each choice presumably made on a random basis, turns out to be as rigidly determined as any other physiological process of the human body.[11]

And without free action, it is further assumed, how can there be any moral responsibility?

Many psychiatrists, to be sure, attempt to escape the hard determinist conclusion that (to quote Hospers) "criminal actions in general are not actions for which their agents are responsible." Yet the escape routes they have proposed end in puzzles that are at least as bad as the puzzle about responsibility that they would escape. Consider four such routes.

The most popular is to take free action as a kind of "first postulate" of both responsibility in law and responsibility in therapy. An example of such a view of therapy is provided by Louisell and Diamond, who assure us that

> it is a mistake to assume that the Freudian psychoanalyst with his emphasis on the specific psychodynamic determinants of behavior abrogates all concept of individual responsibility. Actually, it is an essen-

> tial part of the value system of psychoanalytic therapy that the individual be willing to accept more responsibility for himself and his behavior than society ordinarily assigns. The psychoanalyst insists that the individual must accept responsibility for his own unconscious as well as for his conscious thinking.[12]

An example of such a view of moral and legal responsibility is the view of Alexander and Staub that "we may for practical purposes hold the individual responsible for his acts; that is to say, we assume an attitude as if the conscious Ego actually possessed the power to do what it wishes. Such an attitude has no theoretical foundation, but it has a practical, or still better, a tactical justification."[13]

Lawyers have been very sympathetic with this last "reconciliation" of the deterministic viewpoint of psychiatry with the moral responsibility necessary for law to exact a just punishment. Lawyers are apt to be apologetic and are likely to think that the psychiatrists are right about the inability to do other than one's unconscious causes one to do. In such apologetic moments, lawyers talk of *positing* free human actions while admitting that scientifically there could, of course, be no room for such freedom. For example, Judge Levin, the trial judge in *Pollard*, perceived the deterministic assumptions of the psychiatric testimony before him and responded in this typical way:

> Psychiatry and law approach the problem of human behavior from different philosophical perspectives. Psychiatry purports to be scientific and takes a deterministic position with regard to behavior. "Its view of human nature is expressed in terms of drives and dispositions which, like mechanical forces, operate in accordance with universal laws of causation. . . ." Criminal law is, however, "a practical, rational, normative science which, although it draws upon theoretical science, also is concerned to pass judgment on human conduct. Its views of human nature asserts the reality of free choice and rejects the thesis that the conduct of normal adults is a mere expression of imperious psychological necessity. Given the additional purpose to evaluate conduct, some degree of autonomy is a necessary postulate."[14]

This kind of "reconciliation by fiat" cannot possibly work. The law demands more (and so does therapy if it is to avoid being mere propaganda and indoctrination) than that we can *pretend* that people are free and thus hold them responsible *as if* they were. A just legal system (and a nonpropagandistic therapy) requires that people really *be* responsible. Nor will it do to say that people are free because there is a system of thought (law, morals, or therapy) that makes no sense unless they are free; for the whole point of the hard determinist is that such systems of thought indeed make no sense in light of the scientific truth of determinism. If the hard determinist is right about there being no free action, it

cannot be the case that we can nonetheless found our moral system on some supposed "postulate" of free action *that we believe, as a matter of scientific fact, to be false!* Our moral beliefs cannot be "sealed off" from our scientific beliefs in this way. If our moral beliefs require that we be free, and if determinism shows that we are not, then we cannot be responsible. Attempting to escape that dilemma by simply swallowing a known contradiction is a peculiarly expensive strategy (for what one then gives up is logic itself).

In fairness to the proponents of this view, what they could have in mind is a kind of Kantian dualism about persons: We are "noumenal" beings free of the laws of causation, and we are empirical objects subject to the usual causal laws to which all objects are subject. There is no contradiction, in such a view, because we are free only in one mode of being and determined only in another. The more modern version of this kind of dualism follows the "linguistic turn" of twentieth-century philosophy, talking not about how things are, but rather, about our talk about how things are. According to the more modern "linguistic dualism" there is a category difference between the concepts of intention, choice, and action on the one hand, and the concepts of motion and mechanistic cause, on the other. We have, in this last view, two quite different ways of conceptualizing our selves: as intelligent persons, or as complicated bits of dumb clockwork. Conceptualized as clockwork, we are fully determined; but conceptualized as intelligent agents who make choices, we are free from determinism – free in the sense that it becomes a kind of "category mistake" to say that our actions are caused. Again, there is no contradiction in saying both that we are free and that we are not, because the system of thought in which free is true is different from the system of thought in which determined is true. The completion of this kind of argument would be to say that law and morality view persons in one system of thought (intelligent agents) and psychiatry views persons in the other (a kind of driven clockwork).

The problem with any such attempt to allow us to have both free will and determinism and yet not contradict ourselves, is that there is no categorical barrier to talking of actions being caused. We fully attribute causation of actions to sets of beliefs and desires, when we explain actions by reasons; to mental causes, such as emotions; to environmental stimuli, such as childhood experiences; and to physiological events, when we know them. As we saw in Chapter 4 with regard to Thomas Szasz's extreme use of the doctrine of category difference, there is simply no reason to deny that causal relations exist between Intentionally characterized states, such as actions, and non-Intentionally characterized states, such as physiology.

Even though there are systematic regularities in our usage of the vocabulary of persons that allow us to grant provisional independence

(Chapter 1) to this vocabulary from that of natural science, such categorical differences do not prevent scientific discoveries about what causes our various actions. There is thus no separate system of thought—set of categories—in which we are necessarily free. We cannot "have our cake and eat it too," free of contradiction; we are either free or we are not, but we are not both free and not-free because we and our doings can be described in two different categories.

As unsatisfactory is a second strategy sometimes adopted to avoid the rigors of hard determinism. "Not all psychiatrists are dogmatic in their determinism," Wilber Katz advises us.[15] Such psychiatrists conceive of the possibility of there being "a little bit" of determinism. Sheldon Glueck, for example, finds it helpful to "imagine a simple chart which shows the freedom/determinism proportions of a feeble-minded person, an extreme psychotic, an average 'sociopathic' or psychopathic personality," and others.[16] We might then, Glueck tells us, speculate that the feeble-minded person's freedom/determinism mix "will consist of, say, 10 percent . . . endowed intelligent free-choosing capacity, and 90 percent . . . predetermined blocking of freedom of conscious, purposive choice and control."[17] By contrast, the "chart of the psychopath or sociopath, will consist of, say, 30 percent to 45 percent . . . amount of free choice capacity, the balance rigidly controlled."[18]

Human actions, on this view, can sensibly be said to be more or less determined. Causation, in such a view, is more like baldness than it is like pregnancy: One can be more or less bald (depending on the number of hairs on one's head), but it is difficult to imagine being "a little bit" pregnant.

Such a view, if it could be made out, might be the basis of avoiding a complete collapse of responsibility. For one might hold that only those persons who are "strongly caused" to act are not responsible, but that those who are "weakly caused" to act are responsible. Psychiatrists, in such a case, would separate the responsible from the nonresponsible on the basis of the strength of the unconscious determinants.

Stephen Morse has examined this view in psychiatry recently and has given it the apt name of "selective determinism."[19] It is not, as Morse notes, an exclusively psychiatric view, but has at least partial support in commonsense intuitions about responsibility: "Philosophically impure common sense consistently rejects the philosophically pure view [of the irrelevance of causation to responsibility] by assuming that the behavior of all persons is subject to various causes and that these causes vary in their salience and strength."[20] One might, for example, believe that a slum environment is a "strong and salient" cause of criminal behavior in one person, while believing that an emotionally deprived (but materially well-off) childhood is only a weak cause of criminal behavior in another. Popu-

lar judgment of a liberal sort might distinguish such cases, urging that the second can fairly be held responsible but the first not.[21]

The stunning problem for both psychiatry and common sense here is that there is no sense to the idea of a "little bit" of either causation or of freedom. It makes sense to say that we are determined or that we are free, but to speak of being partly determined or partly free makes as much sense as speaking of being partly pregnant. There are, to be sure, comparative judgments we make about when one cause is more important than another in producing behavior. There is, indeed, quite a literature on the criteria we use in preferring some causally relevant conditions to others in various contexts.[22] But none of this literature can make sense of the quite different comparative judgment we must make here; this is not a comparison of the relative importance of different causes, but rather, a comparison of the importance of *all* causes, on the one hand, and freedom, on the other. How much causation was there? has to be a sensible question to ask, in this view; the problem is that such a question seems to make no sense at all.

Stephen Morse points out that psychiatrists might attempt to work out the needed idea of "degrees of causation" with the idea of *predisposing causation,* that is, causation that only predisposes (makes more likely) bad behavior and that does not operate as a sufficient condition for that behavior.[23] With such a concept, psychiatrists can urge that unconscious forces that are "strongly predisposing" excuse from responsibility without being driven to admitting that no one is responsible. Strongly predisposing here would mean "renders the bad behavior highly likely." It is an essentially probabilistic notion; when certain factors make the probabilities of bad behavior high enough (strongly predisposed), responsibility is said to evaporate.

Morse himself goes on to argue that the unconscious mental state explanations typically proffered by psychiatrists are at most "weakly predisposing"; thus, even by their own criterion psychiatrists have not shown that the unconscious excuses.[24] There is, however, a more fundamental objection here. This is to question whether talking of strongly or weakly predisposing causes can make any sense of the ideas of more causation or of less freedom.

There is no doubt that there are such things as predisposing causes. To give explanations in terms of probabilistic laws – what Carl Hempel calls inductive-statistical explanations[25] – is an unquestioned feature of both natural and social science. It is also undoubtedly true that the probability of bad behavior following various factors varies greatly among such factors; in short, various correlations make bad behavior variously probable. Much more problematic is the assumption, however, that psychiatric explanations are *irreducibly* probabilistic in character.

It may well be that the psychiatric explanations currently available cite factors that only make bad behavior probable; the factors psychiatrists cite as the causes of crime are not sufficient to produce it. Yet this may reflect only our ignorance about *all* the factors truly sufficient to produce the behavior in question. It is a very controversial matter to assert that *there is no such set of sufficient conditions* for human behavior, and that the best one will ever be able to do is to give probabilistic explanations. Imagine a similar claim about the explanations available to explain the pattern of heads/tails landings of a coin that has been flipped repeatedly. The only explanation we might well have for a roughly 50/50 pattern of heads/tails is a probabilistic one. That does not rule out, however, an explanation of each of the coin's landings (and thus of the overall pattern of landings) in terms of a set of sufficient conditions. One might well think that the physical laws governing matter in motion would explain those events if only we knew enough about the original force exerted on the coin each time, about the coin's physical features, the motion of the air, the features of the landing surface, and so forth.

To be sure, in subatomic physics a plausible case has been made for the claim that events are *only* explainable in terms of probabilistic laws. But no similar case has yet been made showing that explanations (psychiatric or otherwise) of the behavior of human beings must be limited to the probabilistic kind. Yet unless such a case is made, the thesis that there are strongly predisposing or weakly predisposing causes is not a metaphysical thesis about there being degrees of freedom from causation, but rather, a thesis having to do with the degree of our present ignorance. Psychiatrists must thus defend the dubious metaphysical position that human behavior, like the behavior of subatomic particles, cannot be explained by sets of sufficient conditions. Those psychiatrists reluctant to defend any such position cannot make sense of the idea of there being degrees of causation via the idea of "predisposing cause."[26] Rather, the latter group of psychiatrists is committed to a robust determinism (with its supposedly dramatic implications for responsibility).

A third purported escape from the rigors of hard determinism is slightly different from that just explored. It is to admit that human behavior, like other events, is fully determined. In this view, it makes no sense to think of there being a "little bit" of causation. Yet one might urge that our responsibility is affected, not by the degree to which we *are* caused to act, but rather by the degree to which we *have knowledge* of such causes.

If studies show us, for example, that 80 percent of kids from a certain environment commit crimes, we would in this view excuse them from responsibility. If there were no such studies, so that we were ignorant of this probabilistic correlation, then those same individuals would be fully responsible. Since our present knowledge about the causes of human behav-

ior is scant, most people would remain responsible. Only where psychiatrists or others could show us strongly predisposing causes would we excuse. In some such way psychiatrists might seek to avoid the abolition of responsibility yet retain the relevance of unconscious mental state causation to it.

Yet it is wildly inconsistent with our other basic beliefs to think that we attribute responsibility based on the fortuity of our present state of knowledge. If we truly believe that behavior is fully determined and that fully caused behavior is not the responsibility of the person whose behavior it is, it is immoral to say that we will nonetheless hold most people responsible because *we* are ignorant of what causes them to behave as they do. *Our* ignorance of a causation that we think universally exists surely can make no difference to another's responsibility. To say otherwise would be like excusing only those who we know have some *particular* excuse, and holding all others responsible because we are ignorant of what excuse they have, even though we believe that all of those others have some valid excuse but we just do not know which one. If causation excuses, it excuses everyone for determinists. There can fairly be no selection based on a fortuity such as how much of the causal story we may happen to know at any particular time.

It may well be that what Morse calls "philosophically impure common sense" here joins psychiatry in urging either the second or the third purported escape route from the rigors of hard determinism. Many people undeniably soften their judgments about responsibility the more they know of the causal story behind any person's bad behavior. *Tout comprendre c'est tout pardonner* does indeed reflect such persons' judgments on responsibility. Yet such philosophically impure common sense should not survive these insights into its philosophical impurity. If common sense believes that there can be a little bit of causation, then it is wrong—wrong because this is inconsistent with our more basic metaphysical ideas of what kinds of causal relations exist. If common sense believes that the degree of a person's responsibility can depend on the degree of our ignorance about a universally present, excusing condition, then common sense is wrong—wrong because this is inconsistent with our more basic moral belief that such fortuitous factors unconnected to the actor have anything to do with his responsibility. Such inconsistencies cannot protect themselves under the mantle of common sense.

As a fourth and last possibility, one might reason in the following way so as to avoid any of the foregoing pitfalls: One would grant that behavior is fully determined, that causation itself, and not our knowledge of it, excuses; but deny that *all* causes excuse. In such a view, only some kinds of causes excuse; if these are not universally present causes of human behavior, then not everyone is excused by determinism.

In their attempt to show that expert testimony about the causation of behavior by unconscious mental states is relevant in criminal trials, Bonnie and Slobogin appear to adopt this view.[27] In order to make out such a view, a morally compelling case must be made to show why some unconscious mental states excuse but others do not. As Stephen Morse has pointed out,[28] totally lacking in the Bonnie and Slobogin view is any such case. They seem to assume that "abnormal intrapsychic forces" of the unconscious excuse,[29] but no case is made for why causation by abnormal forces is an excuse when causation by normal unconscious forces is not. Abnormality itself may preclude responsibility; if abnormal is taken to mean madness, then that is indeed the case. But that in no way advances the argument that unconscious mental state causation excuses. Crazy defendants may be held nonresponsible, but that is because they are crazy, not because they suffer from some peculiar form of causation (Chapter 6).

An alternative possibility would be to urge that unconscious mental state causation always excuses, but to deny that all of our behavior is caused by unconscious mental states (even if it is caused by some set of sufficient factors). This is not, of course, an argument a Freudian psychiatrist can make, given the Freudian's theoretical commitment to a universally operating unconscious that underlies *all* of our conscious thinking and action.

Still, a psychiatrist might claim that Freud was wrong here, and that only some behavior is caused by unconscious mental states. The burden on such a psychiatrist is to show why causation by unconscious mental states excuses when causation by other factors (e.g., the environment, physiology) does not. Prima facie, as we saw in Chapter 3, the pretheoretical unconscious recapturable by extended memory would seem to *expand* responsibility, not contract it.

It might be argued that causation by unconscious mental states is different from the causation of behaviorism or physicalism because the structuralism in psychoanalytic theory posits the interaction of separate agencies "within" each person as the source of unconscious mental state causation. Thus, psychoanalytic causal hypotheses are not about familiar, "dumb" causation; rather, the theory is that we are caused to act by the id, the superego, or the unconscious. We might liken such causation to coercion of the self by our "other selves."

This tack is flatly inconsistent with Freud's assertion that we are responsible for unconsciously governed activities because we are identical with the unconscious id as much as the conscious ego.[30] In any case, this task fails even if one grants the animistic concept of subagencies within the self, because mere causation (not amounting to coercion) of some action by that "other person" is irrelevant to one's responsibility. Suppose one person, X, knows another, Y, to have a limited repertoire of jokes that he will

tell audiences on the proper occasion. Although X can probably *cause* Y to tell the jokes simply by saying something that jogs Y's memory—for example, the first line of one of the jokes—Y's responsibility for telling a bad joke is still his. It is his free, uncoerced action even though caused by another.[31]

With regard to all four of these purported escape routes, the short of it is that psychiatrists are stuck with a genuine dilemma about the determinism of their theory and the responsibility presupposed by both the law and their own therapy. Psychiatrists cannot consistently believe (1) that all behavior is determined by the unconscious; (2) that causation by the unconscious is inconsistent with responsibility; and (3) that persons are morally responsible, not only in morals and law but also in the clinical practice of psychiatry. None of the four strategies we have examined gives any room to psychiatrists to hang on to all three of these beliefs.

In the face of this unavoidable dilemma psychiatrists are prone to eliminating (3), the belief that anyone is really responsible for anything. This is the hard determinist position. In such a case they will think that law and morality make distinctions (between the responsible and the nonresponsible) that are inherently arbitrary. If they participate in the legal system at all, such psychiatrists do so with the sense that they should manipulate whatever archaic formulas the lawyers are these days using, so as to excuse the maximal number of defendants.[32] Lawyers, on the other hand, tend to adopt the strategy of William James, who, believing there to be an irreconcilable conflict between determinism and responsibility, chose to give up determinism (the first belief discussed).[33] Much of the lawyer's talk of law being premised on free will is designed to deny that determinism is applicable to human action.

Neither lawyers nor psychiatrists very often perceive that they need not make such extreme choices. They need not throw out the window that entire sphere of experience we call our moral life, on the grounds that it rests on an illusion; nor need they, in order to validate that moral experience, engage in a bit of speculative metaphysics of a highly dubious sort. The obvious alternative to either of these extreme views is to deny that there is any inconsistency between determinism and responsibility. (One, in other words, gives up the second of the three beliefs just outlined.)

This alternative is not the same as the supposed middle ground explored earlier. For each of those four alternatives assumed that causation was relevant to responsibility, but tried to avoid hard determinism by notions of degrees of causation, inconsistent but nonetheless assertable first postulates, and the like. The present alternative avoids the need of any such notions because causation itself, in any "degree," is simply irrelevant to responsibility. We are fully responsible no matter what causes may exist for our behavior, be they physiological, behavioristic, or psychoanalytic.

To make the case for this position – "soft determinism" or "compatibilism" – is by and large to show the irrelevance of causation to the elements of responsibility.[34] Put another way: It is to show that nowhere in the elements under which responsibility is justly attributed to a person is there any presupposition that the person is *free* (where free means free from being caused to act). There are two elements of responsibility with which causation is commonly confused: action and compulsion. I shall discuss each briefly.

John Hospers's earlier quoted, "puppeteer" view of human beings clearly assumes that causation is relevant to the question of whether there has been an action performed. More specifically, Hospers assumes that causation by external forces precludes there having been an action. With such assumptions, his next step is to liken the unconscious to an external force: "Between an unconscious that willy-nilly determines your actions, and an external force which pushes you, there is little if anything to choose."[35] There *is*, however, something to choose here. In the case of the external force – someone else grabbing X's arm, the wind blowing it, getting hit by an avalanche, or being carried out to sea – X does not act. The notion of action is independent of any notion of causation. As set forth in Chapter 2, a basic act is performed only when the actor knows that he is performing it. Presumably the choice to act, and the action, are always determined by various factors, such as chemical balances in the brain, early environment, or character, but the fact that one acts is completely independent of there being, or not being, causes for the action.

This point is unfortunately confused by certain contemporary philosophical accounts of action that urge that a caused action is inconceivable.[36] These accounts assume that if the wind caused X's arm to move, therefore X did not move it. This is simply a mistake. X did not move his arm because X did not know, consciously or unconsciously, that he was moving it. Whether there was a cause, either external or internal, or whether X knew about it, is irrelevant to whether X moved his arm on some occasion. If the wind caused X's arm to move through pressure, or by causing X to shiver involuntarily, then X did not move his arm. If it caused his arm to move because X, made cold by the wind, put on his jacket, then X did move his arm. Causation by the wind is no guarantee that X does not perform the action of moving his arm. Whether X moves his arm depends on the presence or absence of his nonobservational knowledge, including his memory, that he has moved it.

There is, to be sure, a sense in which the concept of a basic action presupposes a view of persons as autonomous. Yet autonomy here does not mean free of causation. As set forth in Chapter 2, the autonomous person presupposed by our practices in blaming is first and foremost a being with causal powers over his own body. Such autonomy does not

mean that persons must be "uncaused causers." Causal power of a person is not inconsistent with the exercise of those powers being fully determined by other factors. That we are both autonomous agents and fully determined is not a contradiction.

It is true, there can be no *reduction* of the causal agency of a person to event or state causation, as I argued in Chapter 2. But because we cannot *reduce* personal agency to event or state causation does not mean that a person's acts must be uncaused by those states or events. That the concept of basic action does not have as its criterion, causation by certain states of belief or desire, for example, does not rule out the possibility that basic acts are caused by such states of belief or desire. Indeed, as I pointed out earlier, commonly most basic acts are caused by various belief/desire sets, either conscious or unconscious, and they remain basic acts nonetheless.

There is, thus, nothing in the idea of a basic act (nor in the autonomy or causal power a person must have to perform basic acts) that presupposes persons to be free of the laws of causation. The other kind of autonomy presupposed by the concept of a *complex* action also does not require that we view persons as exempt from the laws of the universe. Autonomy in the sense that we have the causal power not only to move our bodies but to do so in such a way as to change external features of the world is similarly uncontaminated by any presuppositions of freedom. Our capacities to reason practically and to perform complex actions is a power we have to change things in the world to suit our wants. Again, the exercise of such capacities on particular occasions can be as determined as one likes; the truth that we have such capacities remains untouched by the truth of determinism. For our basic acts to cause other events to occur in the world does not require that those basic acts themselves be uncaused.

Persons thus need not stand outside the causal order in order to perform human actions. The other element with which causation is commonly confused is compulsion. Compulsion, as we have seen in Chapter 2, is a well-accepted excuse or partial excuse from responsibility. Yet it no more than the concept of action has anything to do with causation. The factors that may quite properly be said to compel us may (often or always) be causes of the behavior they compel. We would properly describe a threat, natural necessity, high emotional states, or addictions as causes. Yet it is not because such factors are causes that they excuse as compulsions. Such factors excuse as compulsions because they make the choice to do what is required very difficult.

It has been a commonplace at least since the writings of Moritz Schlick[37] and A. J. Ayer[38] that causation does not "make us" do certain things in the way in which a gunman makes us do them. Causation analyzed in terms of regularity of sequence is one way of describing the order we find in nature. As Schlick admonishes:

> The laws of nature must not be thought of as supernatural powers forcing nature into a certain behavior . . . but simply as abbreviated expressions of the order in which events do follow each other.[39]

As Schlick elsewhere elaborates:

> Since natural laws are only descriptions of what happens, there can be in regard to them no talk of "compulsion." The laws of celestial mechanics do not prescribe to the planets how they have to move, as though the planets would actually like to move quite otherwise, and are only forced by the burdensome laws of Kepler to move in orderly paths; no, these laws do not in any way "compel" the planets, but express only what in fact planets actually do.[40]

Compulsion, as we saw in Chapter 2, essentially involves interference with practical reasoning. To be compelled is to have one's normal capacities to reason practically interfered with, either by the constraint on means imposed by threats or natural necessity, or by the constraint on ends posed by internal cravings and emotions.

Most causes of behavior do not operate as such constraints. Simply because we are caused to act by our beliefs and desires, for example, can hardly show that we are compelled. That I am caused to go downtown by my desire to get a haircut is hardly a case of compulsion. This is my uncompelled act, the product of my undisturbed practical reasoning.

Similarly, if I am caused to engage in sharp practices by my greedy character, this is not to say I am compelled. My greed is my characteristic way of unconstrainedly dealing with others in financial matters. It does not constrain my powers of practical reasoning so much as describe how I decide when I am unconstrained.[41]

Similarly, there are doubtlessly large numbers of physiological states and events necessary for each of us to engage in various kinds of basic acts. There may even be certain physiological conditions characteristic of what used to be called volitions or acts of will.[42] Such causes hardly disturb our practical reasonings; rather, they are the conditions that make possible the execution of our desires in action.

Consider last the environmental causes that behaviorists tell us can alone be made into sufficient conditions with which to explain adult behavior. Here, one might think, causation *is* compulsion, because with this species of causation our choices are severely constrained. John Hospers, for example, exercises this familiar argument along the way to showing that the unconscious makes none of us responsible:

> Everyone has been moulded by influences which in large measure at least determine his present behavior; he is literally the product of these influences, stemming from periods prior to his "years of discretion," giving him a host of character traits that he cannot change even if he would . . . An act is free when it is determined by the man's character,

> say moralists; but what if the most decisive aspects of his character were already irrevocably acquired before he could do anything to mould them? . . . What are we to say of this kind of "freedom"? Is it not rather like the freedom of the machine to stamp labels on cans when it has been devised for just that purpose?[43]

Hospers's last question is not as rhetorical as he seems to think. The freedom essential to responsibility is the freedom to reason practically without the kind of gross disturbances true compulsions represent. Machines have no such freedom because they have no practical reasoning capacities to begin with and so can hardly be disturbed in the exercise of those capacities. Persons, on the other hand, do have such capacities, and simply because what a person desires or believes is caused by his environment in no way makes the exercise of those capacities difficult.

Causation as such is not the same as compulsion. If one wants to show that some causally relevant factor mitigates responsibility, one can do so only by bringing it within some true moral or legal excuse. A deprived childhood, for example, may cause one to be psychotic, in which event one is not responsible because one's general capacities to reason practically are disturbed. Alternatively, such a childhood may cause one to become addicted to drugs or to have certain compulsive cravings. Such addiction or cravings may operate as a partial excuse because they disturb an agent's ability to do what is required in certain circumstances. In any of such cases, however, it is the disturbance of practical reasoning that excuses, *not* the fact that such disturbance was caused.

The upshot of this is very simple. One can grant as true Freud's entire metapsychological story—about events in childhood combining with instinctual drives to cause adult behavior via the intermediate mechanisms of forces and energies operating in the unconscious—and yet reject any supposed implication for responsibility. That adult behavior is caused in this way is no more relevant to responsibility than are the competing causal stories of behaviorism or physiological psychology. One should thus reject out of hand conclusions such as that of David Rapaport's, that "motivated behaviors are those which we cannot help doing,"[44] reached because (for Rapaport) motivations are drives or drive derivatives and thus are causes of behavior. Such psychoanalytic hard determinism is no more sustainable than is any other form of this philosophically indefensible doctrine.

Given the hard determinism of many psychiatrists, their testimony about the effect of the unconscious on responsibility is of dubious value.[45] The psychiatric report in *Pollard,* for example, read in pertinent part:

> During that period a dissociative state may have existed and his actions may not have been consciously activated.
>
> "It is therefore our opinion that during the period in question, Pollard, while intellectually capable of knowing right from wrong,

> may have been governed by unconscious drives which made it impossible for him to adhere to the right. It is our belief that this unconscious motivation, which could only be positively identified by prolonged analysis, might have been related to guilt feelings in connection with the death of his wife and child, which compelled subsequent acts that would certainly lead to apprehension and punishment.[46]

If the psychiatrists here meant that Pollard's unconscious mental states acted as a compulsion, all well and good. We shall examine shortly when such a case can plausibly be made out. What one suspects, however, is that these psychiatrists drew no distinction between compulsion and causation, that their idea of being compelled is simply to be "governed by unconscious drives," that their idea of impossibility is not that of a hard choice but of a supposed causal necessity. If this latter interpretation of the testimony is correct, then the testimony is irrelevant to any concern of morality or the criminal law.