3

Sadness With and Without Cause

Depression From Ancient Times Through the Nineteenth Century

Depression has been an omnipresent phenomenon over several millennia of human history. For virtually all of that time, from the earliest writings of the ancient Greek physicians to the late twentieth century, Western diagnosticians routinely distinguished depressive disorders, as a form of madness, from symptomatically similar but nondisordered, normal sadness responses to a wide range of painful circumstances. Then, in 1980, seeking a more scientific foundation for diagnosis by focusing on decontextualized criteria based on symptoms, the *DSM-III* inadvertently abandoned this critical traditional distinction, which is now essentially lost in current thinking about depression. This chapter and the next trace the history leading up to this momentous and, we argue, ultimately detrimental conceptual shift.

Why is reviewing this history important? Current diagnostic practices may seem obviously right and sensible just because they are accepted, and they are all that many of us have ever known. To understand the problems with the current diagnostic approach to depressive disorder and to recognize the choices it represents, it helps to place it in historical context. This history reveals that the way we think about depressive disorder now is quite new—and radically diverges from what has traditionally been considered appropriate.

But the importance of history is more than simply providing context and contrast. It is easy to assume that current practices, if they are different, must have emerged from a process in which the traditional alternatives were found to be flawed and were superseded by a superior approach. The history of thinking about depression specifically in regard to the role of context in diagnosing disorder dispels such beliefs and reveals instead the contingency and even arbitrariness of some aspects of current diagnostic practices. It shows that the reasons for the recent divergence from the traditional approach, although well intentioned and shaped by admirable scientific aspirations, are anchored neither in evidence nor

in logic, which in fact support the older tradition. Despite the many virtues of the new approach, it is in certain important respects weaker than those it replaced.

Depressive disorder, unlike many other disorders, has an identifiable and lengthy history. Indeed, depression is probably the psychological disorder that is most easily recognizable throughout history; similar symptomatic descriptions occur over a 2,500-year span, representing what historian Stanley Jackson calls a "remarkable consistency."[1] From the earliest medical texts in ancient Greece to the present *DSM*, deep sadness and its variants—hopelessness, sorrow, dejection, despondency, emptiness, despair, discouragement—were often mentioned as core features of depressive disorder, along with related symptoms such as aversion to food, sleeplessness, irritability, restlessness, feelings of hopelessness or worthlessness, suicidal ideation and attempts, fear of death, repetitive focus on a few negative ideas, lack of pleasure or lack of interest in usual activities, fatigue, and social detachment.

Yet traditional diagnostic treatises also agreed in distinguishing depression as a disorder from a nondisordered type of deep sadness or fear that could have many of the same symptoms but that was a normal, proportionate reaction to serious losses. Such losses included the death of intimates, reversals in fortune, disappointments in attaining valued life goals, romantic disappointments, and the like. In addition, it was traditionally acknowledged that variations in temperament predispose some people to more readily or intensely experience sadness or fear but that these variations could be within a normal range of reasonably proportionate responses that did not represent a disorder.

Depressive disorders differed from these normal reactions, according to tradition, because they either arose in the absence of situations that would normally produce sadness or were of disproportionate magnitude or duration relative to their provoking causes. Such conditions indicated that something was wrong in the individual, not in the environment. In essence, then, traditional psychiatry took a *contextual* approach to the diagnosis of depressive disorder; whether a condition was diagnosed as disordered depended not just on the symptoms, which might be similar in normal sadness, and not just on the condition's severity, for normal sadness can be severe and disordered sadness moderate, but on the degree to which the symptoms were an understandable response to circumstances. In this and the following chapters, we elaborate the history of this contextual approach to depression and how the *DSM-III*, overturning thousands of years of thinking, replaced it with relatively precise and communicable symptomatic criteria that largely ignored the complexities of context, with detrimental side effects for psychiatric diagnosis.

Preliminary Caveats

From ancient Greek medical writings until the early twentieth century, what is now termed *depressive disorder* was generally referred to as *melancholia*, which

literally means "black bile diso
times, it originally reflected the a
the balance or imbalance betwe
excess of black bile—a humor of
responsible for depressive sympt
had a natural function in regul
a failure of this natural functio
waned, *depression* eventually arc
twentieth centuries.

In recognizing the strikingly
order across the millennia, sever
sider the context of each discuss
at all. Like today's confusingly c
and *melancholia* also tended to dc
to normal emotions, moods, and

Second, classic texts were wri
among mental disorders were re
often encompassed what in hind
These included psychotic disorde
and other delusional states. For
scription of the cycling of mania
nize as bipolar disorder may, on c
scription of the alternating agita
who was mistakenly classified as
individuals sometimes have moo
for their sadness, early psychiatri
cholia to others with circumscr
Also, the withdrawal associated
sonality disorder and social pho
for the withdrawal associated wit
ture of those classified as melan
we know it.

Third, because melancholia w
ditions based on their believed ca
placed other conditions that were
bile imbalance together with de
in a broad sense, even if they h
times, these "melancholic disord
Melancholia itself as a disorder
category.

Fourth, classic descriptions g
we would now call *psychotic dep*
tions. Indeed, these descriptions c

literally means "black bile disorder." Although the name stuck into modern times, it originally reflected the ancient belief that health and disease depend on the balance or imbalance between four bodily fluids, or "humors," and that an excess of black bile—a humor often thought to be produced in the spleen—was responsible for depressive symptoms. Ancient physicians thought that black bile had a natural function in regulating mood and that melancholia represented a failure of this natural functioning. As belief in black bile's role in mental life waned, *depression* eventually arose as the dominant term in the nineteenth and twentieth centuries.

In recognizing the strikingly similar clinical descriptions of depressive disorder across the millennia, several cautions are necessary. First, one must consider the context of each discussion to tell whether a disorder is being described at all. Like today's confusingly overused term *depression*, the terms *melancholy* and *melancholia* also tended to do double duty in referring both to a disorder and to normal emotions, moods, and temperaments.

Second, classic texts were written before most of today's refined distinctions among mental disorders were recognized, and thus the category of melancholia often encompassed what in hindsight can be seen to be quite different disorders. These included psychotic disorders that ranged from schizophrenia to paranoid and other delusional states. For example, what may initially appear to be a description of the cycling of mania and depression in what we would now recognize as bipolar disorder may, on closer inspection, turn out to be more likely a description of the alternating agitation and withdrawal of a schizophrenic patient who was mistakenly classified as melancholic.[2] Because psychotically depressed individuals sometimes have mood-congruent delusions that provide the content for their sadness, early psychiatrists sometimes extended the category of melancholia to others with circumscribed delusions that caused negative emotions. Also, the withdrawal associated with such current diagnoses as avoidant personality disorder and social phobia appears to sometimes have been mistaken for the withdrawal associated with melancholia. However, the predominant picture of those classified as melancholics clearly indicates depressive disorder as we know it.

Third, because melancholia was an etiological description that classified conditions based on their believed cause in excess black bile, older descriptions often placed other conditions that were considered to have a similar etiology in black-bile imbalance together with depressive disorders as "melancholic disorders" in a broad sense, even if they had nothing to do with depression. In ancient times, these "melancholic disorders" included, for example, epilepsy and boils. Melancholia itself as a disorder was just one distinct instance of this broader category.

Fourth, classic descriptions generally, though not always, focused on what we would now call *psychotic depression*, which includes delusions or hallucinations. Indeed, these descriptions often defined melancholy as a form of "delirium

without a fever" to distinguish melancholic delusions and hallucinations from those that occurred during a high fever caused by various physical diseases. Melancholia of this kind distinctively involved fixed ideas on specific topics linked to depressive affect, which distinguished melancholia from general cognitive malfunction or psychosis. Nonpsychotic depressive disorders were recognized as well but were not emphasized as constituting the bulk of cases until recent times.

Fifth, there is an ambiguity about the referent of *melancholia* that continues to exist in our own time and that can sometimes cause confusion. The exact extent of the meaning of *melancholia* has varied, sometimes referring to an overall disease, sometimes to sadness as a specific symptom, and sometimes to a syndromal set of coexisting symptoms, of which sadness is just one.[3]

Sixth, contrary to current practice, ancient and many subsequent texts routinely grouped sadness and fear together as symptoms of melancholia. Despondency was thought to be related to fear because melancholics were generally worried or morose not only about actual events but also about negative possibilities in the future that caused apprehension. Contemporary criteria emphasize sadness as an exclusive dominant affect, yet recent studies confirm that anxiety and sadness tend to go together in depression and that it is difficult to distinguish these states, just as tradition would have it.[4] But it is also clear from clinical descriptions that, then as now, sadness alone could be sufficient for melancholia.

A final caveat about our methodology: It is important to look past many differences and confusions in the history of depression to find an underlying coherence and similarity to current judgments. In particular, what are now considered genuine depressive disorders were clearly included within traditional melancholia and were distinguished from normal sadness. No doubt an alternative, postmodernist history of depression might emphasize the social construction of depression, including variations in definitions and in ranges of behavior that were pathologized and the social control correlates of these variations. Although the history of depression certainly contains such elements, the historical record also recognized a common core condition that has been of concern for several millennia.

Indeed, from the earliest times, this record displays what might be considered an "essentialist" view of the classification of melancholia—that is, it involves an inference, common to different writers who may have disagreed in their specific theories, that in melancholia something is going wrong with the internal functioning of mechanisms usually responsible for normal sadness in a way that leads to certain standard symptoms. This view is not an artifact of changing social responses to madness but a considered and plausible judgment. Moreover, it is impossible to analyze responsibly the ways groups have exploited the concept of depression for purposes of social power until one understands the logic of the concept of depression itself. It is the history of this common concept, and especially attempts to distinguish disordered sadness from normal sadness, that we attempt to understand.

The Ancients

Writing in the fifth century B.C
known definition of melancholi
for a long time it is melancholia.
changed, the symptoms that in
and sadness, Hippocrates ment
despondency, sleeplessness, irri
ria.[6] But Hippocrates' definitior
but symptoms of unexpected du
tence that the sadness or fear m
notion that disproportion to cir
disorder.

Indeed, an ancient, possibly
the distinction between disorder
with cause.[7] He was asked to
Macedonia from 454 to 413 B.C
displayed a total lack of concern
the king's condition stemmed fro
deceased father's. He suggested
cubine and secure her love in re
the king suffered not from a mel
ment but from a problem stemmi

A century after Hippocrates,
in the *Problemata* elaborated the
states of sadness on the one han
Aristotle clearly expressed the id
to events. He noted that, if the b
duces groundless despondency."[8]
disproportionate to the circumst
less." Such despondency, for exam
hanging amongst the young and

Aristotle, the master typologis
of melancholy. One distinction
melancholic disorder. In this regar
lasted to our own day of associati
artistic and intellectual ability: "
standing in philosophy, statesman
some to such an extent that they
bile. . . . They are all, as has been
recognized not only melancholic
an abnormal degree of melanchc
may be possessed by. He did not co

delusions and hallucinations from
ed by various physical diseases. Mel-
xed ideas on specific topics linked to
ıcholia from general cognitive mal-
ve disorders were recognized as well
bulk of cases until recent times.
erent of *melancholia* that continues
mes cause confusion. The exact ex-
ı, sometimes referring to an overall
nptom, and sometimes to a syndro-
dness is just one.[3]
nt and many subsequent texts rou-
symptoms of melancholia. Despon-
ause melancholics were generally
ıts but also about negative possibil-
. Contemporary criteria emphasize
recent studies confirm that anxiety
and that it is difficult to distinguish
But it is also clear from clinical de-
ɔuld be sufficient for melancholia.
is important to look past many dif-
pression to find an underlying co-
. In particular, what are now con-
clearly included within traditional
rmal sadness. No doubt an alterna-
ght emphasize the social construc-
finitions and in ranges of behavior
l correlates of these variations. Al-
ɔntains such elements, the histori-
ɔndition that has been of concern

displays what might be considered
nelancholia—that is, it involves an
ıay have disagreed in their specific
ɔing wrong with the internal func-
for normal sadness in a way that
ew is not an artifact of changing
and plausible judgment. Moreover,
ys groups have exploited the con-
er until one understands the logic
tory of this common concept, and
adness from normal sadness, that

The Ancients

Writing in the fifth century B.C., Hippocrates (460–377 B.C.) provided the first known definition of melancholia as a distinct disorder: "If fear or sadness last for a long time it is melancholia."[5] Although theories of depressive disorder have changed, the symptoms that indicate the disorder have not. In addition to fear and sadness, Hippocrates mentioned as possible symptoms "aversion to food, despondency, sleeplessness, irritability, restlessness," much like today's criteria.[6] But Hippocrates' definition indicated that it is not such symptoms alone but symptoms of unexpected duration that indicate disorder. Hippocrates' insistence that the sadness or fear must be prolonged is a first attempt to capture the notion that disproportion to circumstances is an essential aspect of depressive disorder.

Indeed, an ancient, possibly apocryphal, story about Hippocrates illustrates the distinction between disordered sadness without cause and normal sadness with cause.[7] He was asked to diagnose the problem of Perdiccas II, King of Macedonia from 454 to 413 B.C., who had fallen into a morbid condition and displayed a total lack of concern for matters of state. Hippocrates learned that the king's condition stemmed from his secret love for a concubine of his recently deceased father's. He suggested that the king acknowledge his love for the concubine and secure her love in return. In essence, Hippocrates recognized that the king suffered not from a melancholic disease that warranted medical treatment but from a problem stemming from romantic longing.

A century after Hippocrates, Aristotle (384–322 B.C.; or one of his students) in the *Problemata* elaborated the distinction between a variety of normal mood states of sadness on the one hand and pathological disease states on the other. Aristotle clearly expressed the idea that disordered sadness is disproportionate to events. He noted that, if the black bile "be cold beyond due measure, it produces groundless despondency."[8] Here "beyond due measure" refers to what is disproportionate to the circumstances, making the resultant sadness "groundless." Such despondency, for example, "accounts for the prevalence of suicide by hanging amongst the young and sometimes amongst older men too."[9]

Aristotle, the master typologist, suggested several distinctions among types of melancholy. One distinction was between melancholic temperament and melancholic disorder. In this regard, Aristotle inaugurated the tradition that has lasted to our own day of associating depressive temperament with exceptional artistic and intellectual ability: "Why is it that all men who have become outstanding in philosophy, statesmanship, poetry or the arts are melancholic, and some to such an extent that they are infected by the diseases arising from black bile. . . . They are all, as has been said, naturally of this character."[10] Aristotle recognized not only melancholic temperament as a normal variant but also an abnormal degree of melancholy that gifted individuals may possess—and may be possessed by. He did not consider this abnormal degree to be disordered

because it contributed to their creativity, although it did leave them vulnerable to melancholic disorder. "We are often," Aristotle noted, "in the condition of feeling grief without being able to ascribe any cause to it; such feelings occur to a slight degree in everyone, but those who are thoroughly possessed by them acquire them as a permanent part of their nature."[11] In Aristotle's view, such extreme melancholic temperaments were generally disorders except in the rare instances in which they were an integral part of a gifted individual's creativity.

As in Aristotle's passage, the key distinction in ancient definitions of melancholia was between states of sadness *without cause* and those with similar symptoms that arose from actual losses; only the former were mental disorders. But "without cause" did not mean uncaused, for throughout history depression has been attributed to postulated physical or psychological causes such as excessive black bile, disturbances in the circulation of blood, or depletion of energy. Rather, "without cause" meant that the symptoms of depression were not proportional to environmental events that would *appropriately* lead to sadness, such as bereavement, rejection in love, economic failure, and the like.[12] Conversely, ancient Greek and Roman physicians did *not* consider symptoms of depression that occurred "with cause" as signs of a mental disorder because they were normal reactions within their contexts.

Aristotle also grappled with the basic problem of how to define "proportionate" sadness. The puzzle he confronted is this: If the level of sadness or fear varies due to circumstances and has no constant "set point" at which health is defined, how can we define health? Aristotle's answer was that there can be relational definitions of health in which the appropriate amount of sadness varies at any given time in proportion to the circumstances that surround it. That is, Aristotle had the insight to recognize that the relational property of the proportionality between sadness and circumstances can remain present even as the actual amount of sadness and the circumstances vary: "It is possible that even a varying state may be well attempered, and in a sense be a good condition, . . . since the condition may be warmer when necessary and then cold again, or conversely." With such proportionate variation as the baseline, Aristotle went on to conceptualize his notion of the abnormally melancholic—but not strictly disordered—temperament or personality as the tendency of such variation to be extreme on the high end and thus to some degree to overshoot the mark emotionally: "owing to the presence of excess, all melancholic persons are abnormal, not owing to disease but by nature."[13] A similar test would reveal disordered states not due to temperament, as well.

In sum, Aristotle distinguished (1) a melancholic component in all people that gives rise to normal sadness reactions and varying normal moods; (2) a normal-range melancholic temperament in people with a preponderance of black bile and thus an inherent inclination to sadness; (3) an extreme variant of such temperament that often occurs in the gifted and may be considered at least statistically abnormal but is not yet a disease, especially when it is a handmaiden to creativity; and (4) a pathological, harmful, disordered state of disproportionate

sadness or sadness without ade
a creative process. Several of t
in the following passage, which
that in bipolar disorder melan
can be etiologically linked:

> Those who have a small sha
> who have much are unlike
> tense, such men are very m
> kind, they are abnormal. Bu
> ancholic diseases, different p
> the symptoms are epileptic, v
> to deep despondency or to fe

Ancient Roman physicians fol
melancholic states that arose wit
with disease. Thus, for example, t
Hippocrates in defining melanch
fear and sleeplessness"[15] that "
black bile."[16] He advised that, as
should be gently reproved as bein
in the late first or early second ce
cast and prone to anger and . . .
"the signs of melancholy . . . as f
silence, animosity toward membe
and at other times a longing for d
plot is being hatched against him,
ing, and, again, occasional jovial
symptoms.[18] The reference to "w
tion that the emotions of intense

Aretaeus of Cappadocia (ca. A
rion more explicit, noting that m
or unreasonably torpid, without
ment of melancholy. And they a
start up from a disturbed sleep. U
ther distinguish the disordered fr
put it, "mere anger and grief, and
(clearly modeled after the story to
ness that featured symptoms ider
that, consequently, was mistaken

> A story is told, that a certain pe
> girl; and when the physician co

sadness or sadness without adequate cause that is without a redemptive part in a creative process. Several of these distinctions can be discerned, for example, in the following passage, which also interestingly anticipates the modern notion that in bipolar disorder melancholic despondency and manic over-confidence can be etiologically linked:

> Those who have a small share of this temperament are normal, but those who have much are unlike the majority. If the characteristic is very intense, such men are very melancholic, and if the mixture is of a certain kind, they are abnormal. But if they neglect it, they incline towards melancholic diseases, different people in different parts of the body; with some the symptoms are epileptic, with others apoplectic, others again are given to deep despondency or to fear, others are over-confident.[14]

Ancient Roman physicians followed their Greek predecessors in distinguishing melancholic states that arose with and without cause, associating only the latter with disease. Thus, for example, the Roman physician Celsus (ca. A.D. 30) echoed Hippocrates in defining melancholia as "prolonged despondency and prolonged fear and sleeplessness"[15] that "consists in depression which seems caused by black bile."[16] He advised that, as part of the treatment, the patient's "depression should be gently reproved as being without cause."[17] Soranus of Ephesus, writing in the late first or early second century A.D., described the melancholic as "downcast and prone to anger and . . . practically never cheerful and relaxed," with "the signs of melancholy . . . as follows: mental anguish and distress, dejection, silence, animosity toward members of the household, sometimes a desire to live and at other times a longing for death, suspicion on the part of the patient that a plot is being hatched against him, weeping without reason, meaningless muttering, and, again, occasional joviality," as well as various, mostly gastrointestinal, symptoms.[18] The reference to "weeping without reason" makes explicit the notion that the emotions of intense sadness are to some extent without cause.

Aretaeus of Cappadocia (ca. A.D. 150–200) made the "without cause" criterion more explicit, noting that melancholic "patients are dull or stern, dejected or unreasonably torpid, without any manifest cause; such is the commencement of melancholy. And they also become peevish, dispirited, sleepless, and start up from a disturbed sleep. Unreasonable fear also seizes them."[19] To further distinguish the disordered from the normal who experience, as Aretaeus put it, "mere anger and grief, and sad dejection of mind,"[20] he presented a case (clearly modeled after the story told of Hippocrates) of extreme but normal sadness that featured symptoms identical to those occurring in melancholia and that, consequently, was mistaken for a disorder:

> A story is told, that a certain person, incurably affected, fell in love with a girl; and when the physician could bring him no relief, love cured him. But

> I think that he was originally in love, and that he was dejected and spiritless from being unsuccessful with the girl, and appeared to the common people to be melancholic. He then did not know that it was love; but when he imparted the love to the girl, he ceased from his dejection, and dispelled his passion and sorrow; and with joy he awoke from his lowness of spirits, and he became restored to understanding, love being his physician.[21]

Aretaeus thus illustrated how the "without cause" criterion differentiates normal sadness from melancholic disorder, and he pointed to the possibility that normal conditions can be misdiagnosed if symptoms alone are considered.

Like other writers before him, Aretaeus emphasized the delusions of what we would term psychotic depression: "a lowness of spirits from a single phantasy, without fever . . . the understanding is turned . . . in the melancholics to sorrow and despondency only. . . . Those affected with melancholy are not every one of them affected according to one particular form; but they are either suspicious of poisoning, or flee to the desert from misanthropy, or turn superstitious, or contract a hatred of life."[22] Clearly, such delusions, which the literature through to the twentieth century emphasized, provide an alternative to disproportionality as a way of recognizing disorder due to clear cognitive dysfunction.

In the late second century A.D., Claudius Galenus (131–201), known as Galen, like Aretaeus a Greek physician living in Rome, unified and synthesized the psychiatric knowledge that had accumulated over the previous 600 years. Galen simply repeated the Hippocratic definition of melancholia: "Fear or a depressive mood (dysthamia) which lasts for a long time."[23] His description again emphasized psychotic phenomena but also described well the basic symptoms:

> Fear generally befalls the melancholic patients, but the same type of abnormal sensory images do not always present themselves. As for instance, one patient believes he has been turned into a kind of snail and therefore runs away from everyone he meets lest [its shell] should get crushed. . . . Again, another patient is afraid that Atlas who supports the world will become tired and will throw it away and we all will be crushed and pushed together. And there are a thousand other imaginary ideas. . . . Although each melancholic patient acts quite differently than the others, all of them exhibit fear or despondency. They find fault with life and hate people; but not all want to die. For some the fear of death is of principal concern during melancholy. Others again will appear to you quite bizarre because they dread death and desire to die at the same time.[24]

In an implicit acknowledgment of the "without cause" criterion, Galen presented a vivid analogy in which he used the color of black bile to characterize the

fear that the melancholic was
would normally be generated fi

> Because of this despondenc
> constantly sullen and appe
> darkness. As external darkn
> exception of a few naturally
> trained, thus the color of the
> throws a shadow over the ar

The doctrine that emerged
distinguishing melancholic stat
which emotion is "without cau
to external circumstances, pers
regarding melancholia are, how
of Tralles (525–605) included "
of melancholia and recommen
ideas underlying "groundless sa
century Arabic physician Ishaq i
when he defined melancholia pa
tion"; yet he also recognized that
of a beloved child or an irreplace
tion that melancholy is the result
1087) defined melancholia partl
and noted that the loss of a loved
a scholar's loss of his books, coul
emphasized "fear without cause
which do or do not exist; and a
tomarily feared."[30] Often, the "w
explanation that an internal proc
gen's (ca. 1151–1158) descriptic
ter, and releases every evil, somet
in the heart to overflow; it causes
that the person can find no joy in
existence."[31] Not until the Renais
central place it had had in ancien

Depression From the
to the Nineteenth Ce

In the late sixteenth and early
greater emphasis on the "withou

fear that the melancholic was generating from his or her own brain, a fear that would normally be generated from external circumstances:

> Because of this despondency patients hate everyone whom they see, are constantly sullen and appear terrified, like children or adults in deepest darkness. As external darkness renders almost all persons fearful, with the exception of a few naturally audacious ones or those who were specially trained, thus the color of the black humor induces fear when its darkness throws a shadow over the area of thought [in the brain].[25]

The doctrine that emerged in the period between Hippocrates and Galen, distinguishing melancholic states that stemmed from internal dysfunctions in which emotion is "without cause" from those that were proportional reactions to external circumstances, persisted for thousands of years.[26] Explicit sources regarding melancholia are, however, sparse in the following period. Alexander of Tralles (525–605) included "sadness without reason" among the symptoms of melancholia and recommended that, especially in nonchronic cases, the ideas underlying "groundless sadness" should be addressed.[27] The early tenth-century Arabic physician Ishaq ibn Imran reiterated the "without cause" notion when he defined melancholia partly as "irrational, constant sadness and dejection"; yet he also recognized that real losses could trigger true disorder: "The loss of a beloved child or an irreplaceable library can release such sadness and dejection that melancholy is the result."[28] Similarly, Constantinus Africanus (1020?–1087) defined melancholia partly as "fear of things that were not frightening" and noted that the loss of a loved one or of specially beloved possessions, such as a scholar's loss of his books, could trigger melancholia.[29] Avicenna (980–1037) emphasized "fear without cause," including "the appearance of fear of things which do or do not exist; and a greatness of fear of things which are not customarily feared."[30] Often, the "without cause" requirement was implicit in the explanation that an internal process caused the sadness, as in Hildegard of Bingen's (ca. 1151–1158) description: "Melancholy as a Disease. Bile is black, bitter, and releases every evil, sometimes even a brain sickness. It causes the veins in the heart to overflow; it causes depression and doubt in every consolation so that the person can find no joy in heavenly life and no consolation in his earthly existence."[31] Not until the Renaissance, however, did melancholia return to the central place it had had in ancient Greek and Roman psychiatric medicine.

Depression From the Renaissance to the Nineteenth Century

In the late sixteenth and early seventeenth centuries, authors placed even greater emphasis on the "without cause" criterion for disorder. The French

physician Andre Du Laurens (1560–1609), known widely as "Laurentius," wrote *Discourse de la melancholie,* which became known throughout Europe and which heavily influenced later thought. Du Laurens summarized the approach of his time as the "without cause" approach: "A kinde of dotage without any fever, having for his ordinarie companions, feare and sadnes, without any apparent occasion."[32]

On the English side of the Channel, Timothie Bright (1550–1615), a Cambridge-trained doctor of medicine contemporary with Du Laurens, was also much concerned with religious guilt. In his *Treatise of Melancholy* (1568), Bright developed at length the distinction between sorrow with and without cause to allow differential diagnosis between true melancholic disorder and nondisordered states of intense sadness and despair due to the belief that one had sinned and would be the object of God's wrath. He noted that melancholic sadness is such "whereof no occasion was at any time before, nor like to be given hereafter"[33] and argued that "the affliction of soule through conscience of sinne" is "quite another thing than melancholy."[34] Consciousness of sin was "a sorrow and feare upon cause, & that the greatest cause that worketh misery unto man" because of fear of God's wrath, whereas melancholy was "a meere fancy & hath no ground of true and just object." Bright explained in lucid detail how the "Particular difference betwixt melancholy, & the distressed conscience in the same person" which is "the soules proper anguish" could be distinguished based on a contextual understanding of whether the sadness had adequate environmental reasons:

> Whatsoever molestation riseth directly as a proper object of the mind, that in that respect is not melancholicke, but hath a farther ground than fancie, and riseth from conscience, condemning the guylty soule of those ingraven lawes of nature, which no man is voyde of, be he never so barbarous. . . . On the contrarie part, when any conceite troubleth you that hath no sufficient ground of reason, but riseth onely upon the frame of your brayne, which is subject (as hath bene before shewed) unto the humour, that is right melancholicke and so to be accounted of you. These are false points of reason deceaved by the melancholie brayne. . . . Thus I conclude this point of difference, & marke betwixt melancholy and the soules proper anguish. . . . [T]he sense of those that are under this crosse feele an anguish far beyond all affliction of naturall passion, coupled with that organicall feare and heavinesse of heart. The melancholie disposeth to feare, doubt, distrust and heavines, but all either without cause, or where there is cause above it inforceth the passion.

Bright goes on to vividly characterize what the phrase "without cause" means, firmly anchoring the notion in an understanding of the context of the feelings:

> We do see by experience certaine persons which enjoy all the comfortes of this life whatsoever wealth can procure, and whatsoever friendship offereth of kindnes, and wh
> overwhelmed with heavines
> neither receive consolation,
> be neither matter of feare, or
> contrarily of great comfort, a
> by any adversity present or ir

Bright's assumption, which
ancient to modern times, was t
of sadness that was designed to
in disorder.

Subsequent works followed s
Praxeos Medicae (1602) defined
and judgment are so perverted t
sad and fearful. For they cannot
a trivial one or a false opinion wl
apprehension."[36] Like other aut
cause" category both cases lack
lusion or endogenous depressic
which the cause exists but is too

Robert Burton's classic work,
is the most renowned of all Ren
squarely on the "without cause
ponents of depression—mood, c
viewed as the distinguishing feat
melancholic symptoms are not
only symptoms that are without
in this codicil to his definition: "w
all other ordinary passions of F
mind" of melancholia included
ing still, but why they cannot tel

Burton emphasized that a pr
and was a normal and ubiquitou

> Melancholy . . . is either in di
> transitory melancholy which
> of sorrow, need, sickness, tro
> the mind, any manner of care
> guish, dullness, heaviness, an
> ancholy dispositions, no man
> happy, none so patient, so gen
> himself; so well composed, bu
> the smart of it. Melancholy, ir

known widely as "Laurentius,"
e known throughout Europe and
ıurens summarized the approach
"A kinde of dotage without any
feare and sadnes, without any

mothie Bright (1550–1615), a
porary with Du Laurens, was also
atise of Melancholy (1568), Bright
ɔrrow with and without cause to
cholic disorder and nondisordered
ıe belief that one had sinned and
that melancholic sadness is such
or like to be given hereafter"[33] and
science of sinne" is "quite another
in was "a sorrow and feare upon
nisery unto man" because of fear
neere fancy & hath no ground of
detail how the "Particular differ-
cience in the same person" which
uished based on a contextual un-
ate environmental reasons:

a proper object of the mind,
ɛ, but hath a farther ground
ɔndemning the guylty soule of
man is voyde of, be he never
en any conceite troubleth you
ut riseth onely upon the frame
bene before shewed) unto the
ɔ to be accounted of you. These
melancholie brayne. . . . Thus
e betwixt melancholy and the
ıose that are under this crosse
naturall passion, coupled with
ırt. The melancholie disposeth
t all either without cause, or
ɔassion.

ıe phrase "without cause" means,
ng of the context of the feelings:

which enjoy all the comfortes
ɾe, and whatsoever friendship
offereth of kindnes, and whatsoever security may assure them: yet to be overwhelmed with heavines, and dismaide with such feare, as they can neither receive consolation, nor hope of assurance, notwithstanding ther be neither matter of feare, or discontentment, nor yet cause of daunger, but contrarily of great comfort, and gratulation. This passion being not moved by any adversity present or imminent, is attributed to melancholie.[35]

Bright's assumption, which formed the background for the literature from ancient to modern times, was that there exists a "natural passion" or emotion of sadness that was designed to operate a certain way but that had gone wrong in disorder.

Subsequent works followed suit. For example, Felix Platter (1536–1614) in *Praxeos Medicae* (1602) defined melancholy as a state in which "imagination and judgment are so perverted that without any cause the victims become very sad and fearful. For they cannot adduce any certain cause of grief or fear except a trivial one or a false opinion which they have conceived as a result of disturbed apprehension."[36] Like other authors, Platter encompassed within the "without cause" category both cases lacking any actual situational cause (in cases of delusion or endogenous depression) and cases without proportionate cause (in which the cause exists but is too trivial to justify the reaction).

Robert Burton's classic work, *The Anatomy of Melancholy*, published in 1621, is the most renowned of all Renaissance discussions on the topic. It is founded squarely on the "without cause" tradition. Burton described three major components of depression—mood, cognition, and physical symptoms—that are still viewed as the distinguishing features of the condition. However, he insisted that melancholic symptoms are not in themselves sufficient evidence of disorder; only symptoms that are without cause provided such evidence, as he explained in this codicil to his definition: "*without a cause* is lastly inserted, to specify it from all other ordinary passions of Fear and Sorrow." And, he noted, "signs in the mind" of melancholia included "Sorrow . . . without any evident cause; grieving still, but why they cannot tell."[37]

Burton emphasized that a propensity to melancholy was present in all men, and was a normal and ubiquitous aspect of the human condition:

Melancholy . . . is either in disposition or habit. In disposition, it is that transitory melancholy which goes and comes upon every small occasion of sorrow, need, sickness, trouble, fear, grief, passion, or perturbation of the mind, any manner of care, discontent, or thought, which causeth anguish, dullness, heaviness, and vexation of spirit. . . . And from these melancholy dispositions, no man living is free, no Stoic, none so wise, none so happy, none so patient, so generous, so godly, so divine, that can vindicate himself; so well composed, but more or less, some time or other, he feels the smart of it. Melancholy, in this sense is the character of mortality.[38]

In contrast to normal melancholy that arises naturally in people who have suffered loss and disappointment and that is part of the "character of mortality," Burton held that melancholic afflictions are "contrary to nature."[39] This latter condition, the disorder of melancholy, he defined (following Du Laurens) as "a kind of dotage without a fever, having for his ordinary companions fear and sadness, *without any apparent occasion.*"[40]

Burton was sensitive to the wide individual variation in the nature of loss responses, and he allowed a quite broad range of temperamental reactions to loss to be considered nondisordered as long as they did not become chronic and self-perpetuating:

> For that which is but a flea-biting to one, causeth insufferable torment to another, & which one by his singular moderation, & well composed carriage can happily overcome, a second is no whit able to sustaine, but upon every small occasion of misconceived abuse, injurie, griefe, disgrace, losse, crosse, rumor, &c. (if solitary, or idle) yields so farre to passion, that his complexion is altered, his digestion hindred, his sleepe gone, his spirits obscured, and his heart heavy, his Hypocondries misaffected . . . and he himselfe over come with Melancholy. . . . But all these Melancholy fits . . . are but improperly so called, because they continue not; but come & goe, as by some objects they are moved.[41]

It is only when such normal reactions to specific events become established as an ongoing condition independent of events that Burton sees disorder:

> (I)t falleth out oftentimes that these Dispositions become Habits, and . . . make a disease. Even as one Distillation, not yet growne to custome, makes a cough; but continuall and inveterate causeth a consumption of the lungs: so doe these our Melancholy provocations. . . . This Melancholy of which we are to treat . . . a Chronicke or continuate disease, setled humor . . . not errant but fixed . . . growne to an habit, it will hardly be removed.[42]

In addition to noting normal variation in temperament, Burton was an astute observer of the extremes to which normal reactions to loss could go. He noted that the most extremely painful losses included separation from friends and bereavement following loss of a loved one ("in this Labyrinth of accidental causes [of melancholy] . . . loss and death of friends may challenge first place"[43]) and compellingly described the extremes that nondisordered grief can reach:

> If parting of friends, absence alone, can work such violent effects, what shall death do, when they must eternally be separated, never in this world to meet again? This is so grievous a torment for the time, that it takes away their appetite, desire (deep sighs and groans, tear bitter pangs, and by frequen think they see their dead fri still, that good father, that go in their minds; a single thou that are most staid and patie passion of sorrow in this cas times forget themselves, and

It was not only renowned wr practitioners who distinguished cause and those that were propo The work of Richard Napier (15 notebooks have been closely ana lustrates how general physician into three general sorts. The firs row and grief, rejection in love, l spouses, lovers, or parents. Napie adverse states from melancholic fered from the disease of melanc

Two kinds of melancholic stat the term "baseless sorrow" for so cases that were unprovoked or de circumstances. The second type o such as "legitimate occasions in be the sign of melancholy delusic As MacDonald notes, "Contempo melancholy and troubled people The sheer intensity of their moo show that melancholia often aros times stemmed from a disproporti ses of melancholia, for example, loss of a spouse or a child,[49] in v duration that it led to states of ma quired the physician to obtain kno the context of the situations in wh

Writers who followed Burton with and without cause. For exan tury, Timothy Rogers (1658–172 ment as normal response to loss a der. He observed that many peopl "by the loss of Children, by some s

> away their appetite, desire of life, extinguisheth all delights, it causeth deep sighs and groans, tears, exclamations . . . howling, roaring, many bitter pangs, and by frequent mediation extends so far sometimes, they think they see their dead friends continually in their eyes. . . . Still, still, still, that good father, that good son, that good wife, that dear friend runs in their minds; a single thought fills all their mind all year long. . . . They that are most staid and patient are so furiously carried headlong by the passion of sorrow in this case, that brave discreet men otherwise oftentimes forget themselves, and weep like children many months together.[44]

It was not only renowned writers such as Burton but also ordinary medical practitioners who distinguished between melancholic states that arose without cause and those that were proportionate in intensity to their provoking causes. The work of Richard Napier (1559–1634), a physician in rural England whose notebooks have been closely analyzed by the historian Michael MacDonald, illustrates how general physicians of the period classified depressive conditions into three general sorts. The first stemmed from universal experiences of sorrow and grief, rejection in love, loss of fortune, severe illness, and conflicts with spouses, lovers, or parents. Napier explicitly separated these sorts of ubiquitous adverse states from melancholic diseases so that "not every gloomy person suffered from the disease of melancholy."[45]

Two kinds of melancholic states were considered disorders. First, Napier used the term "baseless sorrow" for some of his disordered patients.[46] This referred to cases that were unprovoked or delusional, thus wholly unexplained by external circumstances. The second type of disordered conditions stemmed from sources such as "legitimate occasions in the death of loved ones and were revealed to be the sign of melancholy delusion by their unusual intensity and duration."[47] As MacDonald notes, "Contemporaries believed that the feelings experienced by melancholy and troubled people were exaggerations of normal states of mind. The sheer intensity of their moods was abnormal."[48] Napier's records clearly show that melancholia often arose without situational provocations but sometimes stemmed from a disproportionate response to actual losses. Many diagnoses of melancholia, for example, resulted from bereavement, usually after the loss of a spouse or a child,[49] in which the sadness was of such intensity and duration that it led to states of madness. Judgments of disease consequently required the physician to obtain knowledge of the relationship of the symptoms to the context of the situations in which they arose and persisted.

Writers who followed Burton continued to separate depressions that were with and without cause. For example, toward the end of the seventeenth century, Timothy Rogers (1658–1728) considered the difference between bereavement as normal response to loss and as a triggering cause of depressive disorder. He observed that many people can have a melancholic disorder triggered "by the loss of Children, by some sudden and unlooked for disappointment that

ruines all their former Projects and Designs."[50] But Rogers made it clear that such horrible losses do not usually lead to melancholic disorder. He specifically contrasts such a disordered reaction with that of one Lady Mary Lane, to whom his book is dedicated, who experienced intense but normal grief and sorrow at the loss of her father, mother, and several children.[51]

In the eighteenth century the explicit use of the "without cause" criterion became less common, perhaps because writers in this period focused on psychotic forms of depression in which this description seemed unnecessary.[52] Nevertheless, during this period, madness, according to historian Stanley Jackson, "still usually involved a state of dejection and fearfulness *without an apparent cause,* and some particular circumscribed delusion was still a common feature. Sleeplessness, irritability, restlessness, and constipation continued to be usual elements."[53] Samuel Johnson's famous dictionary, for example, contained three meanings for *melancholia;* two refer to mental disorders and one to common, normal emotions.[54] Incidentally, Johnson was partially responsible for beginning the trend to gradually replace the term *melancholia* with *depression.*

Subsequent medical definitions continued to explicitly use the ancient, contextual definition of melancholia. Friedrich Hoffmann (1660–1742) characterized melancholy as "associated with sadness and fear not having any manifest cause."[55] William Cullen (1710–1790), the preeminent authority on melancholy during the latter part of the eighteenth century, noted that melancholy is "always attended with some seemingly groundless, but very anxious, fear."[56] And in the United States, the famed clergyman Cotton Mather (1663–1728) emphasized the lack of sufficient external justification for sadness in melancholic disorder: "These Melancholicks, do sufficiently *Afflict themselves,* and are Enough their *own Tormentors.* As if this *present Evil World,* would not *Really* afford Sad Things Enough, they create a World of *Imaginary Ones,* and by *Mediating Terror,* they make themselves as Miserable, as they could be from the most *Real Miseries.*"[57] Even the philosopher Immanuel Kant (1724–1804) broadly defined melancholia as "unjustified . . . grief" and carefully distinguished a variety of nondisordered conditions, such as individuals who fashionably immerse themselves in melancholic feelings or the supposed "melancholy mathematician" who in fact is merely introverted and thoughtful, from true mental disorder.[58]

The Nineteenth Century

At the beginning of the nineteenth century, the eminent psychiatrist Philippe Pinel (1745–1826) continued to maintain the fundamental separation between melancholic disorders and the consequences of real misfortunes. In his 1801 book on mental disorder, *Traite Medico-Philosophique Sur l'Alienation Mentale,* Pinel noted that melancholia afflicted "some men otherwise in good health, and frequently in prosperous circumstances. Nothing, however, can be more hideous

than the figure of a melancholi
Pinel also provided a particular
thesis when he distinguished pc
suicide. Observing that the Fren
cross-Channel putdown, distingu
suicides ("the effect of education;
thinking") from disordered Engli
stroy themselves without any ap
and even in the midst of prosper
He elaborates by saying that norm
humiliation or financial reversal,
regarding disordered sources of su
disease: "The propensity to this h
nary powerful motives to it, such a
disease peculiar to England: it is fa

A notable student of Pinel'
1840), continued to embrace the
between reality and the intensity
ferer: "Some . . . possess a knowle
of its falsity, and of the absurdity
perceive clearly that they are irrat
despair."[61] Benjamin Rush (174
psychiatry," similarly allowed for
or disproportionate responses to b

> Partial derangement consists in
> one subject only, with soundne
> jects. The error in this case is t
> it is disproportioned in its effect
> which induce them.[62]

The prominent British psychi
noted the misdirection of the melai
should be agreeable or indifferent
amples of disproportion:

> In some cases it is striking how
> treme mental anguish, the patie
> equate cause for his gloom: one
> very great, said that it was beca
> he ought not to have done, and
> ever because he had muttered a
> prayer.[64]

than the figure of a melancholic, brooding over his imaginary misfortunes."[59] Pinel also provided a particularly important application of the proportionality thesis when he distinguished possible nondisordered from disordered causes of suicide. Observing that the French philosopher Montesquieu, in a sophisticated cross-Channel putdown, distinguished nondisordered culturally shaped Roman suicides ("the effect of education; it depended upon their customs and manner of thinking") from disordered English self-destruction ("The English frequently destroy themselves without any apparent cause to determine them to such an act, and even in the midst of prosperity"), Pinel endorses Montesquieu's distinction. He elaborates by saying that normal triggers for suicide might include severe social humiliation or financial reversal, and he performs an act of medical diplomacy regarding disordered sources of suicide by asserting that this is not just an English disease: "The propensity to this horrid deed as existing independent of the ordinary powerful motives to it, such as the loss of honour or fortune is by no means a disease peculiar to England: it is far from being of rare occurrence in France."[60]

A notable student of Pinel's, Jean-Etienne-Dominique Esquirol (1772–1840), continued to embrace the contextual tradition, noting that the disparity between reality and the intensity of sadness may be apparent even to the sufferer: "Some . . . possess a knowledge of their condition, have a consciousness of its falsity, and of the absurdity of the fears in which they are tormented. They perceive clearly that they are irrational, and often confess it, with grief and even despair."[61] Benjamin Rush (1745–1813), known as "the father of American psychiatry," similarly allowed for melancholia to be characterized by false beliefs or disproportionate responses to beliefs:

> Partial derangement consists in error in opinion, and conduct, upon some one subject only, with soundness of mind upon all, or nearly all other subjects. The error in this case is two-fold. It is directly contrary to truth, or it is disproportioned in its effects, or expected consequences, to the causes which induce them.[62]

The prominent British psychiatrist, Henry Maudsley (1835–1918), also noted the misdirection of the melancholic's response in that "impressions which should be agreeable or indifferent are painful."[63] He offered some extreme examples of disproportion:

> In some cases it is striking how disproportionate the delusion is to the extreme mental anguish, the patient assigning some most ridiculously inadequate cause for his gloom: one man under my care, whose suffering was very great, said that it was because he had drunk a glass of beer which he ought not to have done, and another man was, as he thought, lost for ever because he had muttered a curse when he ought to have uttered a prayer.[64]

Maudsley insisted that any delusional ideas are a result, not a cause, of the affective intensification that comes with the disorder.

The influential German psychiatrist Wilhelm Griesinger (1817–1868) also used the disproportionality of melancholic symptoms to their context to define when they indicated a disorder:

> The melancholia which precedes insanity sometimes appears externally as the direct continuation of some painful emotion dependent upon some objective cause . . . e.g., grief, jealousy; and it is distinguished from the mental pain experienced by healthy persons by its excessive degree, by its more than ordinary protraction, by its becoming more and more independent of external influences, and by the other accessory affections which accompany it. In other cases the melancholia originates without any moral cause.[65]

Greisinger called melancholia "a state of profound emotional perversion, of a depressing and sorrowful character";[66] the intended notion of "perversion" is the turning away of a feeling from the objects at which it would be naturally and proportionately aimed. He noted that melancholia involves the same feelings as in nondisordered responses such as grief and jealousy but that it is distinguished by excessive intensity, duration, and, most of all, its "objective groundlessness" in relation to actual external events.[67] But he acknowledged that "the boundary betwixt the physiological state of emotion and insanity is often difficult to trace" because the disorder "may appear as the immediate continuation of a physiological state of the established emotion." He asserts that "the essential difference" between the disorder of melancholia and a nondisordered "gloomy disposition" is that "in the former the patient cannot withdraw himself from his ill-humour."[68]

Simultaneously with the further elaboration and acceptance of the contextual understanding of depressive disorder, another momentous development was occurring in medical thought. As physicians branched out of the asylum and began to see more patients in private practice, they confronted a much larger proportion of patients coming in for help with intense sadness who had no delusions or other psychotic symptoms. Such forms of melancholia had been recognized since antiquity, but the emphasis had always been on the delusional cases ("dotage without a fever"). But now the form without delusion became singled out as "simple" melancholia, the forerunner of today's nonpsychotic unipolar major depression.

For example, the British psychiatrist D. Hack Tuke (1827–1895) explicitly rejected the idea that melancholia must involve delusion and identified the "simple" form that is purely a matter of symptoms of sadness without cognitive impairment, embracing a category of "melancholia, without delusion" along with a melancholic form of "delusional insanity."[69] He insisted that in simple melancholia there is "no disorde
tion."[70] But nonetheless he detec
sion of a sad, debilitiating, or op
broadly accepted and anticipated
sive disorder that is easiest to con

The greater attention to simpl
focus on contextual criteria in th
ple, psychiatrist John Charles Buc
on diagnosing insanity in a well-
dered symptoms using the "with
delusion:

> The symptoms of melancholia
> existing in a degree far beyond
> ally affect the sane mind, even
> ducing them; and in numerou
> surate moral cause, and often

Due to the nervous system's res
disappears."[73] Bucknill also held t
lia . . . vary in degree, but not in k
sorrow, of which all men have th
precipitating causes, he noted: "it
tal disease; especially by griefs, disa
kind. It is also caused by long-cont
that disorder that was triggered by
disposition as well.[76]

Likewise, psychiatrist Charles M
cholia in Tuke's influential *Dictionc*
on proportionality to actual event
acterized by a feeling of misery wh
cumstances in which the individu
onset until an excessive, dispropor
possible interaction between stress
and the other symptoms reach su
are unmistakably exceeded, and it
ing from a morbid depression."[78] M
intense sadness could be risk factor

> Untoward circumstances, the lo
> ter; any circumstance which is
> easiness, anxiety, in an ordinari
> a person of less than ordinary

melancholia there is "no disorder of the intellect . . . no delusion or hallucination."[70] But nonetheless he detected "a cerebral malady . . . sustained by a passion of a sad, debilitiating, or oppressive character."[71] Such definitions became broadly accepted and anticipated the contemporary focus on the kind of depressive disorder that is easiest to confuse with normal emotional responses.

The greater attention to simple melancholia implied an even more exclusive focus on contextual criteria in the general definition of melancholia. For example, psychiatrist John Charles Bucknill (1817–1897), the author of the chapter on diagnosing insanity in a well-known manual, separated normal from disordered symptoms using the "without cause" criterion but with no reference to delusion:

> The symptoms of melancholia are sorrow, despondency, fear, and despair, existing in a degree far beyond the intensity in which these emotions usually affect the sane mind, even under circumstances most capable of producing them; and in numerous instances existing without any commensurate moral cause, and often without any moral cause whatever.[72]

Due to the nervous system's responses, "proportioned excitement of function disappears."[73] Bucknill also held that symptoms of "uncomplicated melancholia . . . vary in degree, but not in kind, from that normal and healthy grief and sorrow, of which all men have their share in this chequered existence."[74] As to precipitating causes, he noted: "it is occasioned by all the moral causes of mental disease; especially by griefs, disappointments, reverses, and anxieties of every kind. It is also caused by long-continued ill-health."[75] However, Bucknill insisted that disorder that was triggered by normal grief generally required a hereditary disposition as well.[76]

Likewise, psychiatrist Charles Mercier (1852–1918), in his entry on melancholia in Tuke's influential *Dictionary of Psychological Medicine*, relied exclusively on proportionality to actual events in defining melancholia as "a disorder characterized by a feeling of misery which is in excess of what is justified by the circumstances in which the individual is placed."[77] He noted the possible gradual onset until an excessive, disproportionate level of symptoms is reached and the possible interaction between stress and heredity: "At length the degree of misery and the other symptoms reach such a grade at which the limits of the normal are unmistakably exceeded, and it becomes manifest that the patient is suffering from a morbid depression."[78] Mercier recognized that the causes of normal intense sadness could be risk factors for the development of disorder:

> Untoward circumstances, the loss of friends, or of fortune, or of character; any circumstance which is calculated to produce sorrow, grief, uneasiness, anxiety, in an ordinarily constituted person, may, if it acts upon a person of less than ordinary stamina, produce melancholia. . . . The

more severe the stress, the greater, naturally, is the chance of melancholia occurring.[79]

French physician Maurice de Fleury (1860–1931), in *Medicine and the Mind,* characterized the illness simply as "causeless melancholy."[80] He also offered an explanation of how normal grief over time may transform into disorder, analogous to what is these days known as the "kindling hypothesis": "Grief is a special, lower pitch of brain activity. The mind, if it stays there for a certain time, will form the habit, and henceforward everything will appear to it in a painful, melancholy, pessimistic light."[81]

Another psychiatrist, George H. Savage (1842–1921), emphasized the internal causes of melancholic states that were disordered. He defined melancholia as "a state of mental depression, in which the misery is unreasonable either in relation to its apparent cause, or in the peculiar form it assumes, the mental pain depending on physical and bodily changes, and not directly on the *environment.*"[82] Like most other writers, he accepted the category of simple melancholia: "Simple melancholia, i.e., those in whom the misery and its expression are simply slight exaggerations of natural states, those cases in whom there is no real delusion, no fiction such as that they are ruined or damned . . . frequently, the misery gives rise to the delusion."[83]

The most popular psychiatric text of the late nineteenth century, Richard von Krafft-Ebing's (1840–1902) *Text-Book of Insanity,* continued to define melancholia in terms of proportionality of response: "The fundamental phenomenon in melancholia consists of the painful emotional depression, which has no external, or an insufficient external, cause, and general inhibition of the mental activities, which may be entirely arrested."[84]

For Krafft-Ebing:

> A painful, depressed state of feeling . . . that has arisen spontaneously and exists independently, is the fundamental phenomenon in the melancholic states of insanity. . . . Even objects which under other conditions would give rise to pleasant impressions seem now, in the mirror of his abnormally changed sense of self, to be worthy of aversion.[85]

Krafft-Ebing observed the challenge of distinguishing normal from abnormal depressive states, especially in cases of simple melancholia:

> The content of the melancholic consciousness is psychic pain, distress, and depression. . . . This painful depression in its content does not differ from the painful depression due to efficient causes. . . . The content of melancholic delusions is extremely varied, for they include all varieties of human trouble, care, and fear. . . . The common character of all melancholic delusions is that of suffering. . . . Simple melancholia is decidedly

the most frequent form of mer
in institutions for the insane, b
tice (with) innumerable slight

Conclusion

What is striking about this brief
disorder from Hippocrates to Krafft
the symptoms that are mentioned
that current diagnostic manuals e
solid and well-elaborated tradition
tion via various versions of the "v
that goes back to ancient times. Th
derstanding that pathological dep
human emotional response and th
be to use the relation to triggering
disordered. A third point is the re
melancholia" without delusion, yi
"without cause" criterion in definir
disordered sadness and presaging
unipolar disorder. The power, consi
medical understanding of depress
century's radical departures in di
traces the fate of this tradition duri

> the most frequent form of mental disease . . . only exceptionally observed in institutions for the insane, but it is extremely frequent in private practice (with) innumerable slight cases that do not reach the hospital.[86]

Conclusion

What is striking about this brief overview of conceptualizations of depressive disorder from Hippocrates to Krafft-Ebing is, first, the remarkable consistency of the symptoms that are mentioned—by and large the same kinds of symptoms that current diagnostic manuals emphasize. And, second, there is a remarkably solid and well-elaborated tradition of distinguishing disorder from normal emotion via various versions of the "with cause" versus "without cause" criterion that goes back to ancient times. The entire 2,500-year record indicates an understanding that pathological depression is an exaggerated form of a normal human emotional response and thus that the first step in diagnostic logic must be to use the relation to triggering causes to distinguish the normal from the disordered. A third point is the recent move toward greater focus on "simple melancholia" without delusion, yielding even more reliance on the contextual "without cause" criterion in defining the distinction between normal-range and disordered sadness and presaging our contemporary focus on nonpsychotic unipolar disorder. The power, consistency, and rationale of the "without cause" medical understanding of depressive disorder form the backdrop for the next century's radical departures in diagnostic approach. The following chapter traces the fate of this tradition during the twentieth century.

4

Depression in the Twentieth Century

Until the end of the nineteenth century, psychiatry generally used the relationship of symptoms to their provoking causes as an essential part of definitions of melancholic disorder. Although some kinds of cases, such as psychotic depressions, almost always displayed symptoms that implied disorder, diagnosticians understood that they had to consider context, because depressive disorder could often be symptomatically indistinguishable from profound normal sadness. In the late 1800s, the traditional contextual approach to diagnosis of depressive disorder began to divide into two distinct schools. On one side, Sigmund Freud and his followers emphasized the psychological etiology of all mental disorders, including depression, and their continuity with normal functioning. Adherents of this school studied and interpreted the patient's reported thoughts to surmise the existence of underlying unconscious pathogenic meanings and wishes. On the other side, Emil Kraepelin applied a classical medical model that examined the symptoms, course, and prognosis of depression and other disorders to define distinct physical pathologies. Kraepelin's approach inspired a cadre of researchers to translate it into a research program that often used statistical techniques to infer discrete disorders from manifest symptoms.

Many psychiatrists viewed the publication of the *DSM-III* in 1980 as finally resolving the struggle between the Freudian and Kraepelinian schools for the domination of psychiatric nosology largely in favor of Kraepelin's approach.[1] We will see, however, that such a judgment is overly simplistic in many ways. Specifically with respect to depressive disorder, the *DSM-III* criteria in fact represented a rejection of key assumptions underlying both Freud's and Kraepelin's systems and an affirmation of a quite different research tradition that ignored the prior emphasis on contextual criteria.

Continuation of th
Cause Tradition in

Psychodynamic Appro and Normal Sadness

At the beginning of the twent
psychoanalyst Sigmund Freud (
lutionary approach to the study
was the effort to understand pa
mental processes, rather than in
etiologies. Although he acknow
volved in pathogenesis could be
on postulating immediate cause
repressed desires, psychological
motivational energy into anxiet
or other direct physical causes. P
tion to treating symptoms them
underlying, and presumably unc
they thought maintained the syn
and other psychological processe
viewed the psychodynamics tha
tinuous with, not discrete from, t
blurring the boundary between n

For psychoanalysts, depressior
tom formation that, to some degr
postulated a continuum betwee
of depression, and psychotic sta
considered manic depression an e
psychological processes that unde
of self-esteem that all people expe

Analytic attempts to explain d
tions about the differences betwe
without expectable environmenta
ciple of Freud's, provided the firs
grounding his theory in the distin
Abraham considered outwardly si
in fact distinct because they involv
The mourner's grief, Abraham ex
pation with the lost person. In cor
with guilt and low self-esteem. Mo
the depressed person's unconsciou
person; hence the common psycho

tieth Century

century, psychiatry generally used
to their provoking causes as an
disorder. Although some kinds of
t always displayed symptoms that
that they had to consider context,
ymptomatically indistinguishable
1800s, the traditional contextual
ler began to divide into two dis-
and his followers emphasized the
s, including depression, and their
ents of this school studied and in-
surmise the existence of under-
d wishes. On the other side, Emil
el that examined the symptoms,
r disorders to define distinct physi-
ed a cadre of researchers to trans-
ised statistical techniques to infer

n of the *DSM-III* in 1980 as finally
and Kraepelinian schools for the
in favor of Kraepelin's approach.[1]
is overly simplistic in many ways.
r, the *DSM-III* criteria in fact repre-
lying both Freud's and Kraepelin's
nt research tradition that ignored

Continuation of the "With" and "Without" Cause Tradition in the Twentieth Century

Psychodynamic Approaches to Disordered and Normal Sadness

At the beginning of the twentieth century, the Austrian neurologist-turned-psychoanalyst Sigmund Freud (1856–1939) and his disciples developed a revolutionary approach to the study of mental disorders. The heart of this approach was the effort to understand pathological symptoms in terms of unconscious mental processes, rather than in terms of biological predispositions and organic etiologies. Although he acknowledged that the intensity of specific desires involved in pathogenesis could be indirectly due to constitution, Freud focused on postulating immediate causes that were often purely psychogenic, such as repressed desires, psychological conflicts, or the transformation of repressed motivational energy into anxiety, all of which had little to do with hereditary or other direct physical causes. Psychoanalysts paid relatively little direct attention to treating symptoms themselves and focused instead on identifying the underlying, and presumably unconscious, dynamics of mental disorders, which they thought maintained the symptoms. In addition, given the sorts of conflicts and other psychological processes they postulated as etiologies, psychoanalysts viewed the psychodynamics that underlie mental disorders as generally continuous with, not discrete from, the psychodynamics present in normality, thus blurring the boundary between normality and disorder.

For psychoanalysts, depression was one major mechanism underlying symptom formation that, to some degree, was present in nearly every neurosis. They postulated a continuum between ordinary states of sadness, neurotic states of depression, and psychotic states of melancholia. Analysts, for example, considered manic depression an extremely exaggerated expression of the same psychological processes that underlie the universal heightening and reduction of self-esteem that all people experience.[2]

Analytic attempts to explain depression were based on traditional assumptions about the differences between depressive conditions that arose with and without expectable environmental causes. Karl Abraham (1877–1925), a disciple of Freud's, provided the first psychoanalytic explanation of depression, grounding his theory in the distinction between normal grief and depression.[3] Abraham considered outwardly similar states, such as grief and depression, as in fact distinct because they involved different underlying etiological dynamics. The mourner's grief, Abraham explained, stemmed from a conscious preoccupation with the lost person. In contrast, the depressed person was preoccupied with guilt and low self-esteem. Moreover, symptoms of depression resulted from the depressed person's unconscious turning inward of hostility toward another person; hence the common psychoanalytic description of depression as "anger

turned inward" and resultant therapeutic strategies aimed at having the patient express the repressed anger.

Freud elaborated on Abraham's distinction between normal grief and depression in his central article on depression, "Mourning and Melancholia." Freud began his essay by noting the differences between normal grief and melancholia and explaining that

> Although grief involves grave departures from the normal attitude to life, it never occurs to us to regard it as a morbid condition and hand the mourner over to medical treatment. We rest assured that after a lapse of time it will be overcome, and we look upon any interference with it as inadvisable or even harmful.[4]

Freud distinguished between the normality of grief and the disorder of melancholia. He asserted that symptoms associated with mourning are intense and are "grave departures from the normal," in the sense that grief is greatly different from usual functioning. Nevertheless, grief is not a "morbid" condition; that is, it is not a medical disorder that represents the breakdown of a biologically normal response. Thus it does not require medical treatment; indeed, Freud emphasized that it would "never occur to us" to provide medical treatment to the bereaved. In addition, he stressed that grief is naturally self-healing, so that with time the mourner would return to a normal psychological state. Medical intervention, he suggested, could actually harm the grieving person through interfering with this natural process.

While noting that mourners did not suffer from the same unwarranted decline in self-esteem that characterized melancholics, Freud emphasized that their symptoms were otherwise similar. Both mourning and melancholia featured profound dejection, loss of interest in the outside world, an inability to feel pleasure, and an inhibition of activity. The distinction between mourning and melancholia lay not so much in their symptoms but in the fact that the former state was a normal reaction to loss, whereas the latter state was pathological.

Freud's version of the distinction between depressions with cause (mourning) and without cause (melancholia) allowed him to elucidate the different psychodynamics that underlay the two conditions. For mourners, the world came to feel empty and without meaning due to conscious losses, whereas melancholics experienced the ego as impoverished due to unconscious losses. The self-reproaches of melancholics pathologically redirected their internalized hostility from earlier love objects onto the self. Therapy, therefore, should teach them to express their inward anger toward the objects that are its actual targets. In contrast, people who experience normal sadness are going through a natural and necessary process that it was "inadvisable or even harmful" to disrupt with medical treatment.

Freud rejected the 2,500-year
of pathological depression and a
Nonetheless, Freud and other psyc
traditional distinction between no
symptomatically quite similar patl

Kraepelin and Depressiv

Emil Kraepelin (1856–1926), a
Freud, attempted to place psycl
work that considered mental dis
pathologies. He used the symptom
that, he claimed, represented distin
he hoped would eventually be con
sions. He built on earlier work tha
those who might be restored to the
Kraepelin famously used prognosis
sanity (now bipolar disorder), whic
dementia praecox (now schizophr
course, as two fundamental forms

Kraepelin's contributions to ps
categorization based on careful att
as the forerunner of the later *DSM*
Indeed, the recent *DSM*s are now of
prominent historians of medicine,
DSM-III approach, see Kraepelin as
passing even Freud: "It is Kraepelin
the central figure in the history of
approach has become linked to tha
views at some length.

Kraepelin began his career as a
tained his almost exclusive interest
delberg and as the director of the I
his classification system using desc
institutions had become a common
during the nineteenth century.[8] Bef
as those of Richard Napier, would
who treated a great variety of sever
sadness stemmed from life problem
themselves, sought help from friend
or clergy.[9]

The effect of the mental hospital
turbed, and only this group, within a

Freud rejected the 2,500-year tradition that postulated physiological causes of pathological depression and adopted a psychogenic theory of causation. Nonetheless, Freud and other psychoanalysts largely accepted as self-evident the traditional distinction between normal intense sadness resulting from loss and symptomatically quite similar pathological depression disproportionate to loss.

Kraepelin and Depressive Disorder

Emil Kraepelin (1856–1926), a German psychiatrist and a contemporary of Freud, attempted to place psychiatry within a strictly biomedical framework that considered mental disorders as manifestations of physical brain pathologies. He used the symptoms and course of disorders to create categories that, he claimed, represented distinct underlying pathological conditions, which he hoped would eventually be confirmed by the identification of anatomical lesions. He built on earlier work that attempted to separate asylum patients into those who might be restored to the community versus those likely to deteriorate. Kraepelin famously used prognosis to distinguish between manic-depressive insanity (now bipolar disorder), which tended to occur in episodes and remit, and dementia praecox (now schizophrenia), which tended to have a deteriorating course, as two fundamental forms of psychotic disorder.

Kraepelin's contributions to psychiatric diagnosis, especially his efforts at categorization based on careful attention to symptoms, are now generally seen as the forerunner of the later *DSM-III* transformation of psychiatric diagnosis. Indeed, the recent *DSM*s are now often referred to as "neo-Kraepelinian."[5] Some prominent historians of medicine, prompted by his perceived relationship to the *DSM-III* approach, see Kraepelin as the major figure in modern psychiatry, surpassing even Freud: "It is Kraepelin," asserts Edward Shorter, "not Freud, who is the central figure in the history of psychiatry."[6] Because Kraepelin's diagnostic approach has become linked to that of the *DSM*, it is pertinent to consider his views at some length.

Kraepelin began his career as a physician in a Munich asylum and maintained his almost exclusive interest in psychotic disorders as a professor at Heidelberg and as the director of the Psychiatric Clinic at Munich.[7] He developed his classification system using descriptions of inpatient cases. Inpatient mental institutions had become a common setting for treating the seriously mentally ill during the nineteenth century.[8] Before this time, most depressed patients, such as those of Richard Napier, would have visited community-based physicians who treated a great variety of severe and less severe conditions. Persons whose sadness stemmed from life problems would have typically handled the problem themselves, sought help from friends and family, or consulted general physicians or clergy.[9]

The effect of the mental hospital was to concentrate the most seriously disturbed, and only this group, within a single location. Those who entered asylums

would typically have had such severe conditions that the issue of whether or not their current symptoms were proportionate responses to their circumstances would not have arisen. The pressing question for Kraepelin, therefore, was not whether asylum patients had disorders or normal unhappiness but rather what particular types of disorders they had.

Kraepelin confronted a field in intellectual chaos, with no consensual diagnostic system. Everyone since Greek times had used symptoms to individuate disorders. But without any commonly shared principle for how to divide up the varied symptomatic presentations that physicians and psychiatrists saw, the use of symptoms allowed for many different classification schemes. At one extreme were those who classified virtually any symptom presentation as a separate disorder, leading to disorder proliferation that could reach hundreds of categories. At the other extreme were those who, focusing on psychosis, considered all mental disorders to be variants of a single disorder.[10] For example, the first U.S. census survey to ask about mental disorder in 1840 reflected the latter approach and contained just one category of mental disorder, "insanity."[11]

Kraepelin's careful attention to symptoms and their course in inferring distinct pathological states that caused the symptoms followed a tradition in physical medicine started by the eighteenth-century English physician Thomas Sydenham and developed by the nineteenth-century German pathologist Rudolph Virchow. This approach had been highly successful in helping to distinguish physical diseases, especially as knowledge of infectious agents and physical pathology rapidly grew.[12]

Kraepelin was no doubt also greatly influenced by the growing realization that one of the most dreadful mental disorders of his time, general paresis (about which he wrote a book), resulted from the syphilitic infection of the nervous system. This startling discovery seemed to impart two lessons. First, mental disorders, like physical disorders, could be due to underlying physical pathology of some kind and thus fit directly within traditional diagnostic theory. Second, diagnosticians identified general paresis as a specific syndrome based on its symptoms and its horrific and rapid course and poor prognosis; like syphilis itself, the symptoms changed over time and could differ markedly at different stages of the disease, yet the same underlying disorder was present and simply unfolding. The moral seemed clear; it is not just symptoms at any particular time but symptoms over the course of an illness that served to identify the illness.

Kraepelin's descriptions of the depressive symptoms that occur in the course of various affective or mood disorders—which included psychic symptoms, such as slowness of thinking, sense of hopelessness, inner torment, inhibited activity, and inability to feel pleasure, as well as physical symptoms, such as sleep and appetite disorders and fatigue—remain the basis of current diagnostic classifications of depressive disorders. A cornerstone of Kraepelin's thinking was that a great variety of symptomatic presentations of affective disorders in fact

represented one underlying pat
unity of various symptom preser
only depressed and had no mani
der. "In the course of the years,
and more convinced that all (me
of a *single morbid process.*"[13] Kra
represented variations of the sa
depressive states was based on th
the frequent appearance of mar
course of disorders that initially d
many affective patients had depr
Kraepelin also included within
mood disorders that pass "withou
predisposition," under the assun
ments of and often developed into

Kraepelin also maintained tha
reditary predispositions; consequ
may be to an astonishing degree *ir*
cases that seemed to arise norma
quarrels, unrequited love, infidelit
festations of disorders that stemm
Kraepelin wrote, "of the malady n
which at least very often, perhaps
be distinguished from normality b
inexplicable recurrence, psychotic
tion of the trigger.

The relationship between Krae
complex and less clear than is oft
DSM-III, psychiatrist Robert Spitze
grounds that he assumes neither th
gies that underlie different syndrom
to physical brain diseases, both ba
fundamentally, Kraepelin rejected t
necessary and sufficient indicators
evidence, including the prognosis o
tions were likely due to the same pat
against the sole use of symptomatic
ent. Of course, diagnosticians have t
Kraepelin did this in a way that was
underlying pathology, an approach i
on operational definitions solely via

Kraepelin's approach to diagn
pended on the prior identification o

represented one underlying pathology. Based on this hypothetical underlying unity of various symptom presentations, he classified even individuals who were only depressed and had no manic symptoms as having manic-depressive disorder. "In the course of the years," Kraepelin emphasized, "I have become more and more convinced that all (melancholic) states only represent manifestations of a *single morbid process.*"[13] Kraepelin's belief that unipolar depressive states represented variations of the same underlying illness condition as did manic-depressive states was based on the evidence of their overlapping symptoms and the frequent appearance of manic symptoms during recurrences later in the course of disorders that initially displayed only depressive symptoms. Over time, many affective patients had depressive states, manic states, and mixed states. Kraepelin also included within the manic-depressive category even "slight" mood disorders that pass "without sharp boundary into the domain of personal predisposition," under the assumption that these mild conditions were rudiments of and often developed into more severe disorders.[14]

Kraepelin also maintained that most affective disorders stemmed from hereditary predispositions; consequently, "attacks of manic-depressive insanity may be to an astonishing degree *independent of external influences.*"[15] Even many cases that seemed to arise normally from external influences such as deaths, quarrels, unrequited love, infidelity, or financial difficulties actually were manifestations of disorders that stemmed from innate dispositions. "The real cause," Kraepelin wrote, "of the malady must be sought in *permanent internal changes,* which at least very often, perhaps always, are innate."[16] These conditions could be distinguished from normality by telltale evidence such as manic symptoms, inexplicable recurrence, psychotic ideation, or duration well beyond the cessation of the trigger.

The relationship between Kraepelin's work and the *DSM-III* revolution is complex and less clear than is often maintained. The major developer of the *DSM-III,* psychiatrist Robert Spitzer, denies being a "neo-Kraepelinian" on the grounds that he assumes neither that there must be distinct categorical pathologies that underlie different syndromes nor that mental disorders are largely due to physical brain diseases, both basic tenets of Kraepelin's approach.[17] Most fundamentally, Kraepelin rejected the use of any rigid system of symptoms as necessary and sufficient indicators of disorder. Instead, he used all the available evidence, including the prognosis of symptoms, to infer whether various conditions were likely due to the same pathology. He was, contrary to common belief, *against* the sole use of symptomatic criteria to infer which disorder was present. Of course, diagnosticians have to use symptoms as their main resource, but Kraepelin did this in a way that was intended to transcend symptoms and get at underlying pathology, an approach in tension with the *DSM-III*'s heavy reliance on operational definitions solely via symptom syndromes.

Kraepelin's approach to diagnosing distinct pathologies obviously depended on the prior identification of conditions as pathologies, distinct from

nonpathological states that do not involve any underlying pathological etiology. How, then, did Kraepelin deal with the distinction between normal sadness and disorder?

Kraepelin and Normal Sadness

Previous commentators have not examined Kraepelin's approach to distinguishing normal sadness from disorder. Admittedly, his works contain little explicitly about this distinction. As noted, the asylum context in which he worked tended to make this distinction irrelevant, as all his patients likely had disorders. Moreover, Kraepelin, like many psychiatrists, was more worried about false negatives and the harm that missing a true case could do than about false positives that mislabel a normal person as disordered.

Nonetheless, Kraepelin required such a distinction, and he embraced the same doctrine as had the medical tradition that preceded him, namely, that nondisordered intense sadness occurs in response to a variety of losses and can symptomatically resemble depressive disorder. Kraepelin thus accepted the traditional principle that the way to distinguish pathological depressive disorder from normal sadness was to determine whether the sadness was without cause (or without proportional cause). Although he did not explicitly state the "without cause" principle directly in his diagnostic criteria, he did make his position on normal sadness clear in scattered remarks:

> Morbid emotions are distinguished from healthy emotions chiefly through the lack of a sufficient cause, as well as by their intensity and persistence. . . . Even in normal life moods come and go in an unaccountable way, but we are always able to control and dispel them, while morbid moods defy all attempts at control. Again, morbid emotions sometimes attach themselves to some certain external occasions, but they do not vanish with the cause like normal feelings, and they acquire a certain independence.[18]

Here, Kraepelin emphasized that either morbid states were without "sufficient cause" in circumstances or, when they initially seem to be with cause, they became independent of circumstances and continued even after circumstances changed. Such cases include conditions that were initially disorders, as well as conditions that began as normal responses but subsequently became morbid.

Kraepelin addressed the issue of differentiating between disorder and normal sadness in some of his case presentations, such as the following:

> I will first place before you a farmer, aged fifty-nine, who was admitted to the hospital a year ago. . . . On being questioned about his illness, he breaks into lamentations, saying that he did not tell the whole truth on his admission, but concealed the f
> and practiced uncleanness wi
> "I am so apprehensive, so wret
> I had only not transgressed so g
> seven or eight months before hi
> Loss of appetite and dyspepsia
> The most striking feature of th
> *sion.* At first sight, it resembles
> patient says that he was alway
> worse. But there is not the leas
> yet it has lasted for months, wit
> sign of its morbidity.[19]

Kraepelin noted that even the
toms of this patient were consist
in a person with a dispositional
he observed, the patient's sympto
Moreover, in addition to the fact t
the apprehension," the condition h
and seemingly inordinate duratio
do, displayed a trajectory of decrea
creasing severity" over time even t
stances to warrant such changes.
from external events, and especial
coping and mastery, "is the diagno

Kraepelin diagnosed this patien
would surely also qualify for the *D*
on the basis of the duration and s
ing sleep problems, appetite proble
guilt and self-reproach. But Kraep
this depressive disorder from norma
not in the case of this patient but in
semble this patient's in manifest syn
discussion is that, after reciting the
"at first sight, it resembles the anxie
somewhat melancholic (but norma
lamentations and guilt remind one
in the last chapter, of cases of inte
sinned against God's law.) That is, h
duration and severity can be a nor
Kraepelin, not the duration or symp
tional relation to any plausible exte
condition was a disorder. In contrast

admission, but concealed the fact that he had fallen into sin in his youth and practiced uncleanness with himself; everything he did was wrong. "I am so apprehensive, so wretched; I cannot lie still for anxiety. O God, if I had only not transgressed so grievously!" . . . The illness began gradually seven or eight months before his admission, without any assignable cause. Loss of appetite and dyspepsia appeared first, and then ideas of sin. . . . The most striking feature of this clinical picture is the *apprehensive depression.* At first sight, it resembles the anxieties of a healthy person, and the patient says that he was always rather apprehensive, and has only grown worse. But there is not the least external cause for the apprehension, and yet it has lasted for months, with increasing severity. This is the diagnostic sign of its morbidity.[19]

Kraepelin noted that even the extreme emotional and physiological symptoms of this patient were consistent with intense normal sadness, especially in a person with a dispositional tendency toward the melancholic side. But, he observed, the patient's symptoms started "without any assignable cause." Moreover, in addition to the fact that "there is not the least external cause for the apprehension," the condition had lasted months (and thus has a prolonged and seemingly inordinate duration) and had not, as normal sadness episodes do, displayed a trajectory of decreasing symptoms; far from it, it has shown "increasing severity" over time even though nothing new occurred in the circumstances to warrant such changes. This disconnection of the patient's condition from external events, and especially the lack of a trajectory showing normal coping and mastery, "is the diagnostic sign of its morbidity."

Kraepelin diagnosed this patient as depressively disordered, and the patient would surely also qualify for the *DSM* diagnosis of Major Depressive Disorder on the basis of the duration and symptoms of the depressive episode, including sleep problems, appetite problems, depressive mood, and intense unjustified guilt and self-reproach. But Kraepelin's comments on differential diagnosis of this depressive disorder from normal sadness imply a divergence from the *DSM,* not in the case of this patient but in the cases of normal responses that might resemble this patient's in manifest symptomatology. What is critical in Kraepelin's discussion is that, after reciting the duration and the symptoms, he noted that "at first sight, it resembles the anxieties of a healthy person," especially one with somewhat melancholic (but normal range) temperament. (Indeed, the patient's lamentations and guilt remind one of Timothie Bright's descriptions, reviewed in the last chapter, of cases of intense normal guilt due to believing one has sinned against God's law.) That is, Kraepelin recognized that symptoms of this duration and severity can be a normal response to events. It is, according to Kraepelin, not the duration or symptoms in themselves but their lack of proportional relation to any plausible external cause that allowed him to see that this condition was a disorder. In contrast, based on its symptom and 2-week duration

criteria, the *DSM* would automatically diagnose such an individual as depressively disordered without the kind of assessment Kraepelin performed. From the *DSM*'s perspective, Kraepelin's painstaking discussion is pointless because the possibility of normal response does not exist given the symptoms, so there is no differential diagnosis to be made.

In another passage in which he reiterated the "without cause" criterion as central to diagnosis, Kraepelin made it clear that, even in his day when more severe cases were the rule among psychiatric patients, there was a real possibility of misclassifying a normal person as disordered because the symptoms could be identical:

> Under certain circumstances it may become very difficult to distinguish an attack of manic-depressive insanity from a *psychogenic* state of depression. Several times patients have been brought to me, whose deep dejection, poverty of expression, and anxious tension tempt to the assumption of a circular depression, while it came out afterwards, that they were cases of moodiness, which had for their cause serious delinquencies and threatened legal proceedings. As the slighter depressions of manic-depressive insanity, as far as we are able to make a survey, may wholly resemble the well-founded moodiness of health, with the essential difference that they arise without occasion, it will sometimes not be possible straightway to arrive at a correct interpretation without knowledge of the previous history in cases of the kind mentioned.[20]

Although Kraepelin recognized some psychogenic depressions (i.e., those caused by strictly psychological factors that do not include whatever biological pathology underlies manic-depressive conditions) as disorders, he also used the term *psychogenic* to refer to normal sadness states with sufficient external cause. The crucial point, which Kraepelin derived from his experience, is that "the slighter depressions of manic-depressive insanity, as far as we are able to make a survey, may wholly resemble the well-founded moodiness of health, with the essential difference that they arise without occasion."

Kraepelin acknowledged that he initially believed that the patients in question were disordered, noting that the facts about the context that reversed his judgment only "came out afterwards." This confirms that Kraepelin understood that the symptomatic presentation of normal and disordered cases could be the same, and it explains why he emphasized that the causal context was the essential differentiating criterion. It is also worth noting that none of the normal cases he reported encountering involved bereavement, the one contextual consideration the *DSM* allows, but rather "had for their cause serious delinquencies and threatened legal proceedings." Thus, as we shall see, the *DSM* would likely classify as disordered these cases that Kraepelin diagnosed as normal because it ignores the "essential difference" of context.

Consider another of Kraepelin
between disorder and normality:

> I will now show you a widow,
> efforts to take her own life. Th
> ried at the age of thirty, and h
> husband died two years ago,
> obliged to sell her home at th
> be divided, she grew apprehe
> to want, although, on quiet c
> groundless. . . . This patient, toc
> gives connected information a
> sions, apart from fear that she
> that the real meaning of the w
> *apprehensive depression,* with the
> tal agitation in the sane—i.e., lc
> general nutrition. The resembla
> greater because the depression
> we can easily see that the severit
> emotional depression have gone
> patient herself sees clearly enou
> by her real position in life, and t
> should wish to die.[21]

This patient was experiencing h
toms; there was "no insane history."
the patient was suicidal and had in
("failure of the general nutrition")
of MDD. Although the depressive
death, the immediate trigger seems
need to sell their home and attenda
the financial and social consequenc
mal reaction. Once again, the sympt
occur in nondisordered people who
accompaniments as we see in menta
blance to anxiety in the sane person
followed a painful external cause."

How, then, did Kraepelin know t
there was a trigger, the reaction, wh
serious suicide attempts, went beyo
to the trigger: "the severity, and mor
depression have gone beyond the lim
that the feelings were without caus

Consider another of Kraepelin's cases that raises the issue of the distinction between disorder and normality:

> I will now show you a widow, aged fifty-four, who has made very serious efforts to take her own life. This patient has no insane history. She married at the age of thirty, and has four healthy children. She says that her husband died two years ago, and since then she has slept badly. Being obliged to sell her home at that time, because the inheritance was to be divided, she grew apprehensive, and thought that she would come to want, although, on quiet consideration, she saw that her fears were groundless. . . . This patient, too, is quite clear as to her surroundings, and gives connected information about her condition. She has no real delusions, apart from fear that she will never be well again. Indeed, we find that the real meaning of the whole picture of disease is only permanent *apprehensive depression*, with the same accompaniments as we see in mental agitation in the sane—i.e., loss of sleep and appetite, and failure of the general nutrition. The resemblance to anxiety in the sane person is all the greater because the depression has followed a painful external cause. But we can easily see that the severity, and more especially the duration, of the emotional depression have gone beyond the limits of what is normal. The patient herself sees clearly enough that her apprehension is not justified by her real position in life, and that there is absolutely no reason why she should wish to die.[21]

This patient was experiencing her one and only episode of depressive symptoms; there was "no insane history." In addition to manifesting depressed mood, the patient was suicidal and had insomnia, loss of appetite, and lack of energy ("failure of the general nutrition") and so would qualify for a *DSM* diagnosis of MDD. Although the depressive symptoms began soon after her husband's death, the immediate trigger seems to have been not that but the subsequent need to sell their home and attendant fears of poverty; as we saw in chapter 2, the financial and social consequences of loss can influence the severity of a normal reaction. Once again, the symptoms—including even suicidality, which can occur in nondisordered people who are highly distraught—consist of "the same accompaniments as we see in mental agitation in the sane." Indeed, "the resemblance to anxiety in the sane person is all the greater because the depression has followed a painful external cause."

How, then, did Kraepelin know that this woman was disordered? Although there was a trigger, the reaction, which had lasted about 2 years and included serious suicide attempts, went beyond any possible proportional relationship to the trigger: "the severity, and more especially the duration, of the emotional depression have gone beyond the limits of what is normal." In effect, this meant that the feelings were without cause. This was apparent even to the patient

herself: "On quiet consideration, she saw that her fears were groundless. . . . The patient herself sees clearly enough that her apprehension is not justified by her real position in life, and that there is absolutely no reason why she should wish to die." Indeed, the patient had every reason to live, including four healthy children. The case illustrates that when the severity and duration of symptoms are disproportionate to the trigger, they are in effect symptoms "without cause" because the context interacting with normal human nature does not fully explain them. As Kraepelin elsewhere emphasized, "The dejection which in normal life accompanies sad experiences gradually wanes, but in disease even a cheerful environment fails to mitigate sadness, indeed, it may even intensify it."[22]

In sum, Kraepelin maintained the traditional distinction between depressive conditions that were "with" or "without cause." Not symptoms in themselves, but symptoms that became detached from their contexts and took on a life of their own, indicated disorder. Kraepelin offered symptoms as evidence to infer a diagnosis but, in contrast to the *DSM*, he never attempted to define disorders solely in terms of necessary and sufficient symptoms. He clearly recognized normal depressive episodes "with cause" that were proportionate to their triggers and that subsided after the stressor subsided, and he actively grappled with how to distinguish normal sadness from disorder given their possible symptomatic similarity.

Adolf Meyer on Normal and Disordered Reaction Types

Adolf Meyer (1866–1950), a Swiss-born psychiatrist who held the Chair in Psychiatry at Johns Hopkins University, is generally considered the leading American psychiatrist in the first half of the twentieth century. Both the Kraepelinian physiological and Freudian psychological traditions influenced Meyer, and he was known early on for bringing Kraepelinian ideas to American psychiatry, but he was not a full-fledged partisan of either school. By the 1920s, he developed his own distinctive approach that focused more on life course, personality, and patients' capacity for responding to adaptive challenges and less on the particular diseases they might have. Indeed, he reconceptualized psychiatric disorders as impairments in the ability to respond to such everyday problems. Meyer's approach heavily influenced the descriptions of disorders in the first two editions of the *DSM* that preceded the pivotal third edition.

Like psychoanalysts, Meyer emphasized a contextual approach to depression. He thought that the symptoms, causes, and prognoses of depressive illnesses were far too heterogeneous to be encompassed within a single disease condition. Instead, he developed a "biopsychosocial" approach, which stressed how each individual's unique predispositions, environmental circumstances, and specific experiences over the life course produced their conditions. For Meyer, psychiatric disorders, including depression, were maladaptive reactions that

arose on the basis of constitutio
ual upbringing, and social cond
vidual organisms and their envir
"without cause" tradition, Meye
and altogether unjustified depres
excesses of normal depression."[23]

In response to Kraepelin's foc
underlying physical pathologies
Meyer included a constitutional (b
ceiving of mental pathologies as r
ity to react adaptively to stressful
for thinking about all disorders th
reaction, and final adjustment."[24]
in psychopathology are more or l
reactions and adjustments, Meye
to loss in his conception of patholo
as malfunctioning responses to eve
understand disordered individuals
mental contexts.

In principle, Meyer and his follo
distinction between normal reacti
reactions that were excessive and
that the distinction between norma
toms but in the relation to events.
chobiology and Psychiatry (1939),
disorder as a reaction that is differe
mal sadness via its disproportionali

> Depression is a sweeping reacti
> of sadness or its equivalent ap
> syndrome. . . . The mood may be
> choly, or more topically pointed
> The reaction presents general s
> loss of initiative . . . slowness in
> self-depreciation, etc. Pathologic
> normal depression by its greater
> the causative factors. Depression
> ated since depression of normal

Note that Muncie implicitly assu
sion, although proportionate to ca
logical depressions. Indeed, he wait
complete to add criteria for distin

arose on the basis of constitutional and psychological predispositions, individual upbringing, and social conditions, as well as from the interaction of individual organisms and their environments. In a definition that accorded with the "without cause" tradition, Meyer defined simple melancholia as "an excessive and altogether unjustified depression" and simple depression as "more or less, excesses of normal depression."[23]

In response to Kraepelin's focus on classification as involving inference to underlying physical pathologies analogous to diagnosis of physical diseases, Meyer included a constitutional (biological) component to stress reactions. Conceiving of mental pathologies as malfunctions in the individual's overall capacity to react adaptively to stressful situations, he developed a general framework for thinking about all disorders that was summarized in the schema, "situation, reaction, and final adjustment."[24] Meyer argued that "the conditions we meet in psychopathology are more or less abnormal reaction types."[25] In talking of reactions and adjustments, Meyer did not include normal sadness in response to loss in his conception of pathology. Rather, conceiving of episodes of disorder as malfunctioning responses to events, he was essentially urging psychiatrists to understand disordered individuals as reacting dysfunctionally to their environmental contexts.

In principle, Meyer and his followers held to a clear, coherent, and traditional distinction between normal reactions that were proportionate and disordered reactions that were excessive and disproportionate. They also clearly discerned that the distinction between normal and disordered depression lies not in symptoms but in the relation to events. Wendell Muncie's Meyerian textbook, *Psychobiology and Psychiatry* (1939), with a foreword by Meyer, defined depressive disorder as a reaction that is differentiated from the universal experience of normal sadness via its disproportionality:

> Depression is a sweeping reaction in which a dominant and fixed mood of sadness or its equivalent appears as the central issue determining a syndrome. . . . The mood may be rather diffuse as sadness, blueness, melancholy, or more topically pointed as worry, or fearful or anxious depression. The reaction presents general slowing and reduction of useful activity, loss of initiative . . . slowness in thinking . . . ideas of unworthiness, and self-depreciation, etc. Pathological depression is to be differentiated from normal depression by its greater fixity, depth, and by the disproportion to the causative factors. Depression is the major reaction most easily appreciated since depression of normal proportions is a universal experience.[26]

Note that Muncie implicitly assumed that the symptoms of normal depression, although proportionate to causes, were similar to those of some pathological depressions. Indeed, he waited until after his symptom description was complete to add criteria for distinguishing the two kinds of depression via

the familiar, classic criteria: greater duration ("fixity"), unusual severity of symptoms ("depth"), and disproportion to the cause.

Both Meyer and psychoanalysts focused their concern more on understanding personalities and life circumstances than on distinguishing distinct disease conditions. Their greatest classificatory impact was on the diagnostic manuals that preceded the *DSM-III*, the *DSM-I* and *DSM-II*, which adopted Meyer's "reaction" vocabulary and psychoanalytic ideas about anxiety and defense in some of their definitions, including the definition of depressive disorder.

Initial Psychiatric Classifications

Psychiatric nomenclature in the United States during the first half of the twentieth century did not reflect an intense interest in classification. Instead, the administrative need to keep track of statistics regarding disorders in groups such as hospitalized patients drove the development of diagnostic manuals.[27] Diagnoses focused on the conditions of people found within institutional contexts, the predominant form of treatment of mental disorder at the time, and reflected the fact that most psychiatrists practiced in mental hospitals. Thus diagnostic systems tended to gloss over the less severe neurotic conditions that analysts typically saw in outpatient settings. For example, the first standardized classification system in the United States, the *Statistical Manual for the Use of Hospitals for Mental Diseases*, issued in 1918, divided mental disorders into 22 principal groups, only one of which represented all psychoneuroses.[28]

The *Statistical Manual* contains two categories that covered depressive conditions. First, one of the 2 groups was for non-neurotic disorders of psychogenic origin without clearly defined hereditary or constitutional causes. Manic-depressive psychoses fell into this category (in sharp distinction to Kraepelin's biological view and more akin to psychodynamic approaches). Second, under the general group of *psychoneurosis*, was the category of depression under the label *reactive depression*, in a Meyerian spirit. Its definition of reactive depression is:

> Here are to be classified those cases which show depression in reaction to obvious external causes which might naturally produce sadness, such as bereavement, sickness and financial and other worries. The reaction, of a more marked degree and of longer duration than normal sadness, may be looked upon as pathological. The deep depression, with motor and mental retardation, shown in the manic-depressive depression is not present, but these reactions may be more closely related in fact to the manic-depressive reactions than to the psychoneuroses.[29]

This definition recognized that depressive disorder is to be distinguished from sadness that arises proportionally "with cause" from external circumstances, which is produced "naturally" (i.e., in accord with human nature) and

thus is normal and not patholog
recognizing that a broad range c
sadness, offering a clearly nonexh
cal illness, and financial reversals
ger recognize the range of potenti

The *Statistical Manual*'s distin
pression, not so different from tha
tinction but required pathological
duration ("of a more marked deg
ness"). They were not of the depth
they were still disorders. An exam
not determine pathology, which w
were of disproportionate intensity
hat to Kraepelin, speculated that
an underlying etiological factor v
plaining their unwarranted intens
mirrored the same three kinds of c
cause, and of disproportionate seve
Robert Burton delineated in *Anato*
recognized that only the latter two

The *Statistical Manual* guided ps
1918 through its 10th edition in 19
in American psychiatry had shiftec
chotic cases, to psychodynamic out
classifications of psychotic disorder
thus no longer relevant to the vast r
Psychiatric Association newly codi
edition of a new manual, the *Diagno*
(DSM-I),[30] that better reflected the r
ing patient population.

A combination of psychodynam
characterization of depression in th
logical aspects of disorders and focu
nisms.[31] It contained one category o
were divided into manic-depressive
tions. Both of these conditions show
evidence of gross misinterpretation
hallucinations."[32] The former also fe
ject to remission and recurrence, w
swings but frequently featured envir

The manual characterized psyc
psychoneuroses, as stemming from
ety, a basically psychoanalytic pers

thus is normal and not pathological. The definition also followed tradition in recognizing that a broad range of negative circumstances can trigger normal sadness, offering a clearly nonexhaustive list of examples, including grief, medical illness, and financial reversals, in contrast to recent definitions that no longer recognize the range of potential triggers of intense normal sadness.

The *Statistical Manual*'s distinction between normal and pathological depression, not so different from that of Hippocrates, offered no symptomatic distinction but required pathological depressions to be more severe and of longer duration ("of a more marked degree and of longer duration than normal sadness"). They were not of the depth and severity of manic-depressive illness, yet they were still disorders. An examination of symptoms alone, therefore, could not determine pathology, which was recognized to exist only when symptoms were of disproportionate intensity to their context. The *Manual*, in a tip of the hat to Kraepelin, speculated that pathological depressive reactions may share an underlying etiological factor with manic-depressive depressions, thus explaining their unwarranted intensity and disproportion. Indeed, this definition mirrored the same three kinds of conditions—depressions with cause, without cause, and of disproportionate severity and duration to a provoking cause—that Robert Burton delineated in *Anatomy of Melancholy;* like Burton's definition, it recognized that only the latter two conditions indicated mental disorder.

The *Statistical Manual* guided psychiatric classification from its 1st edition in 1918 through its 10th edition in 1942. By the early 1950s, the center of gravity in American psychiatry had shifted from state hospitals, which focused on psychotic cases, to psychodynamic outpatient therapy of less severe pathology. The classifications of psychotic disorders that dominated the *Statistical Manual* were thus no longer relevant to the vast majority of patients. In 1952, the American Psychiatric Association newly codified mental disorders and produced the first edition of a new manual, the *Diagnostic and Statistical Manual of Mental Disorders (DSM-I)*,[30] that better reflected the nature of the psychiatric profession's changing patient population.

A combination of psychodynamic and Meyerian approaches dominated the characterization of depression in the *DSM-I*, which generally downplayed biological aspects of disorders and focused on unconscious psychological mechanisms.[31] It contained one category of psychotic affective reactions that, in turn, were divided into manic-depressive reactions and psychotic-depressive reactions. Both of these conditions showed severe symptoms that involved "manifest evidence of gross misinterpretation of reality, including, at times, delusions and hallucinations."[32] The former also featured severe mood swings that were subject to remission and recurrence, whereas the latter did not encompass mood swings but frequently featured environmental precipitating factors.

The manual characterized psychoneurotic depressive disorders, like all psychoneuroses, as stemming from unconscious attempts to deal with anxiety, a basically psychoanalytic perspective. Again in a Meyerian fashion, as

a variation of the earlier *Statistical Manual*'s "reactive depressions," *DSM-I* labeled these conditions *Depressive reactions*, which it defined as follows:

> The anxiety in this reaction is allayed, and hence partially relieved, by depression and self-depreciation. The reaction is precipitated by a current situation, frequently by some loss sustained by the patient, and is often associated with a feeling of guilt for past failures or deeds. The degree of the reaction in such cases is dependent upon the intensity of the patient's ambivalent feeling toward his loss (love, possession) as well as upon the realistic circumstances of the loss.
>
> The term is synonymous with "reactive depression" and is to be differentiated from the corresponding psychotic reaction. In this differentiation, points to be considered are (1) life history of patient, with special reference to mood swings (suggestive of psychotic reaction), to the personality structure (neurotic or cyclothymic) and to precipitating environmental factors and (2) absence of malignant symptoms (hypochondriacal preoccupation, agitation, delusions, particularly somatic, hallucinations, severe guilt feelings, intractable insomnia, suicidal ruminations, severe psychomotor retardation, profound retardation of thought, stupor).[33]

This definition of depressive reactions relied heavily on psychodynamic speculations about etiology to define depressive neuroses. The *DSM-I* not only conceived of depressive conditions as ways that people attempt to defend against underlying states of anxiety but also infused the definition of depression with dynamic assumptions that guilt and feelings of ambivalence were central components of the condition. Aside from such etiological defining criteria, much of the definition was taken up with distinguishing psychoneurotic-depressive disorders from psychotic-depressive disorders.

The *DSM-I*'s definition of depressive reaction might appear to be a historical anomaly in that it did not say a word about the distinction between disordered psychoneurotic-depressive reactions and normal reactions to circumstances. This lapse was more apparent than real, however, because the distinction was implicit, based in the *DSM-I*'s underlying psychodynamic etiological assumptions. Spelling out the distinction between normal and disordered depressive responses was superfluous precisely because the *DSM-I* relied on a theory of etiology to identify disorders and to distinguish them, by implication, from normal conditions in which the etiology is absent. The definition, in effect, specified the dysfunctions of psychological mechanisms that caused the intensity of the sadness, including unwarranted guilt and self-deprecation, intense ambivalence about the lost object, and the use of defense mechanisms (including depressive feelings) to avoid the natural anxieties that arise from loss situations. These processes combined to lead to a depressive response that was not merely sadness that was proportional to any actual loss itself (although the "current situation" and the "realistic circumstances of the loss

was, rather, an inflated, dispropor

of these internal psychological dy

gers of loss responses that might b

the loss, disordered—are loss of lo

The *DSM-I* was the official ma

successor, the *DSM-II*, provides a

neurosis," as follows:

> This disorder is manifested by
> an internal conflict or to an id
> object or cherished possession.
> melancholia and Manic-depres
> sive reactions are to be classifie

The *DSM-II* implicitly recognize

proportionate responses to loss and

tionate. The definition assumed that

depression and attempted neither to

symptoms to distinguish disorder fr

etiology, in the form of internal co

definition also recognized that loss

reaction even in the absence of inte

triggers beyond the loss of a loved c

some extent, the *DSM-II* definition

specifying disordered depression as

In sum, 2,500 years of psychiat

a propensity to potentially intense

can be judged to exist, it was wide

of triggering events fail to establish

symptoms. The major influences on

twentieth century—Freud, Kraepeli

als, such as *DSM-I* and *II*, that they i

explicitly or implicitly embraced this

The Breakdown of the Cause Tradition

The Post-Kraepelinians

During the half-century between a

psychodynamic views of Freud and

"realistic circumstances of the loss" influenced the response's intensity) but that was, rather, an inflated, disproportionate "degree of reaction" due to the action of these internal psychological dysfunctions. Note that the examples of the triggers of loss responses that might be normal—or, if there were ambivalence about the loss, disordered—are loss of love and of possessions, not bereavement.

The *DSM-I* was the official manual of the APA between 1952 and 1968. Its successor, the *DSM-II*, provides a much more succinct definition of "depressive neurosis," as follows:

> This disorder is manifested by an excessive reaction of depression due to an internal conflict or to an identifiable event such as the loss of a love object or cherished possession. It is to be distinguished from Involutional melancholia and Manic-depressive illness. Reactive depressions or depressive reactions are to be classified here.[34]

The *DSM-II* implicitly recognized the distinction between depressions that were proportionate responses to loss and those that were "excessive" and thus disproportionate. The definition assumed that psychiatrists knew what symptoms constituted depression and attempted neither to specify them nor to suggest that one could use symptoms to distinguish disorder from nondisorder. Again, the definition relied on etiology, in the form of internal conflict, to suggest internal dysfunction, but the definition also recognized that losses may trigger a disproportionate, disordered reaction even in the absence of internal conflict. The definition also noted normal triggers beyond the loss of a loved one, such as loss of a cherished possession. To some extent, the *DSM-II* definition was a return to the classic tradition of simply specifying disordered depression as a disproportionate, "excessive" response.

In sum, 2,500 years of psychiatry held that normal human nature included a propensity to potentially intense sadness after certain kinds of losses. Disorder can be judged to exist, it was widely agreed, only when explanations in terms of triggering events fail to establish a normal cause for the intensity or duration of symptoms. The major influences on psychiatric classification in the first half of the twentieth century—Freud, Kraepelin, and Meyer and the early diagnostic manuals, such as *DSM-I* and *II*, that they influenced—disagreed on many things, but all explicitly or implicitly embraced this understanding of depressive disorder.

The Breakdown of the "With" and "Without" Cause Tradition

The Post-Kraepelinians

During the half-century between about 1920 and 1970, the dominance of the psychodynamic views of Freud and the context-based views of Meyer ensured

the general neglect of Kraepelin's system of categorization, which assumed underlying physical etiologies. However, Kraepelin's approach inspired some researchers, especially in the United Kingdom, to pursue an extensive agenda of research into classification of types of depression.

Numerous empirical studies examined symptom patterns in an attempt to discover whether depression consisted of one or more distinct disorders. The work of the psychiatrist Aubrey Lewis was especially influential. In 1934 Lewis published a study of 61 patients treated at the Maudsley Hospital in London.[35] He argued that the distinction between endogenous and reactive depressions was untenable because most supposedly endogenous depressions had external precipitating factors; also, a lifetime of dispositions to depression preceded most reactive depressions. Lewis's research seemed to confirm Kraepelin's claim that almost all depression is one disorder, varying along a continuum of severity from mild to severe but not differing by endogenous or reactive causes. A few researchers, confirming Lewis's contentions, found that depressive symptoms were continuous, and they could not discern patterns that were sufficiently robust to suggest differing underlying etiologies. This group, like Kraepelin, concluded that a rigid division between endogenous and reactive or neurotic and psychotic depressions was unjustified.[36]

Most researchers, however, rejected the notion that all forms of depression fell on a single continuum. Instead, they found that *endogenous* or *psychotic* depression appeared to be a distinct type. The symptoms of psychotic depressions, which often featured hallucinations and delusions, did not correlate with the symptoms of other types of depressions and showed distinct responses to treatment;[37] psychotic depressions seemed more responsive to both electroconvulsive treatment and the antidepressant drug imipramine and less responsive to placebo treatments than other depressed states.[38] Efforts to distinguish psychotic depressions by their lack of environmental precipitants, however, were usually not successful.[39] Instead, stressful life events usually preceded the emergence of all sorts of depressions. Given the paucity of nontriggered depressions that were truly "without cause," the term *endogenous* gradually came to refer to a phenomenological pattern of symptoms, not to a particular cause of symptoms. *Psychotic* or *severe* more accurately characterized the nature of this condition.

Although researchers in this period generally came to agree that psychotic (or endogenous) depressions constituted one distinct type of depression, they could not agree on the nature of nonpsychotic depressions. Gradually, the use of *neurotic* prevailed over *reactive* because precipitating events in the environment provoked the great majority of all types of depression. Some concluded that depression was binary, featuring a neurotic type, as well as a psychotic one.[40] Others felt that three or more distinct types of neurotic depressions existed, although they differed on both the number and the nature of these states.[41] In contrast to the relatively homogenous symptoms found in psychotic depressions, neurotic symptoms were heterogeneous and diffuse across studies.[42] Depending

on the study, neurotic depression
reflected helplessness, low self-es
tility, irritation, and disappointm
schemes.

For our purposes of understar
tailed content and substantive re
gram are not as significant as its m
about the nature of depression fro
terns between 1920 and 1970, th
quent revolution in psychiatric dia
took to identifying depressive diso
Kraepelin, but their approach in fa
during this period relied on measu
point in time. Researchers largely
pecially, the situational context of
rejected using symptoms in thems
sion and emphasized instead the n
conditions as well as the importanc
ordered sadness on the basis of con

The symptom-based emphasis
newly developed statistical meth
whether depression was a single ill
sis attempts to distinguish various
to which individual symptoms ten
There is no inherent, in-principle
and the consideration of the propo
ness of emotional reactions as pa
tual practice, however, the compl
researchers to deviate from the cli
terns alone, without regard to eith
different types of depression. Base
they studied were often hospitalize
dered and had already been diagn
techniques to isolate symptom pa
that all the symptoms they entered
disorder in the sampled populations
criteria that eventually emerged fro
be applied far beyond the clearly dis
derived to progressively broader gr
not mean the same thing.

Lewis's finding that most depressi
made the decision to focus on sympt
haps context in the form of "with c

f categorization, which assumed
pelin's approach inspired some re-
to pursue an extensive agenda of
ion.
mptom patterns in an attempt to
e or more distinct disorders. The
ecially influential. In 1934 Lewis
e Maudsley Hospital in London.[35]
ogenous and reactive depressions
ogenous depressions had external
tions to depression preceded most
to confirm Kraepelin's claim that
g along a continuum of severity
genous or reactive causes. A few
found that depressive symptoms
n patterns that were sufficiently
s. This group, like Kraepelin, con-
ous and reactive or neurotic and

otion that all forms of depression
d that *endogenous* or *psychotic* de-
mptoms of psychotic depressions,
sions, did not correlate with the
howed distinct responses to treat-
sponsive to both electroconvulsive
amine and less responsive to pla-
[38] Efforts to distinguish psychotic
ecipitants, however, were usually
sually preceded the emergence of
ontriggered depressions that were
gradually came to refer to a phe-
a particular cause of symptoms.
zed the nature of this condition.
ally came to agree that psychotic
distinct type of depression, they
depressions. Gradually, the use of
itating events in the environment
lepression. Some concluded that
ype, as well as a psychotic one.[40]
f neurotic depressions existed, al-
d the nature of these states.[41] In
ns found in psychotic depressions,
liffuse across studies.[42] Depending

on the study, neurotic depression featured some combination of symptoms that reflected helplessness, low self-esteem, dysphoria, demoralization, anger, hostility, irritation, and disappointment reactions that resisted precise diagnostic schemes.

For our purposes of understanding the roots of current diagnosis, the detailed content and substantive results of this post-Kraepelinian research program are not as significant as its methodology. Although no consensus emerged about the nature of depression from empirical research regarding symptom patterns between 1920 and 1970, the studies did help pave the way for the subsequent revolution in psychiatric diagnoses because of the general approach they took to identifying depressive disorder. These researchers claimed to emulate Kraepelin, but their approach in fact sharply diverged from his. Empirical studies during this period relied on measuring only symptom presentations at a single point in time. Researchers largely set aside issues of course, duration, and, especially, the situational context of symptoms. In contrast, as we saw, Kraepelin rejected using symptoms in themselves to distinguish varying types of depression and emphasized instead the need to examine the course and prognosis of conditions as well as the importance of distinguishing between normal and disordered sadness on the basis of context.

The symptom-based emphasis reflected the way researchers had exploited newly developed statistical methods, especially factor analysis, to analyze whether depression was a single illness or had multiple types.[43] Factor analysis attempts to distinguish various symptom clusters by examining the extent to which individual symptoms tend to occur together with other symptoms. There is no inherent, in-principle conflict between such statistical methods and the consideration of the proportionality of symptoms or the reasonableness of emotional reactions as part of what is statistically analyzed. In actual practice, however, the complexity that such judgments introduced led researchers to deviate from the clinical tradition and rely on symptom patterns alone, without regard to either their context or course, to distinguish different types of depression. Based on the fact that the clinical populations they studied were often hospitalized and in any event generally clearly disordered and had already been diagnosed, researchers who relied on statistical techniques to isolate symptom patterns simply assumed, quite reasonably, that all the symptoms they entered into their models were manifestations of disorder in the sampled populations. But, as we shall see, the kinds of clinical criteria that eventually emerged from these symptom-based analyses came to be applied far beyond the clearly disordered populations from which they were derived to progressively broader groups in which the same symptoms might not mean the same thing.

Lewis's finding that most depressions followed some kind of triggering event made the decision to focus on symptoms easier, because it suggested that perhaps context in the form of "with cause" versus "without cause" was not so

important after all.[44] However, Lewis's research never explored the notion of the disproportionality of a response to the nature of the reported trigger, which was at the heart of the classic tradition. Moreover, Lewis's study was of an inpatient, clearly disordered, sample, so it could not reveal differences between the disordered and the nondisordered.

The replacement of the "with cause" or "without cause" distinction by categories based on types of symptoms had especially dire consequences for misdiagnosis of normal individuals because of a major change in the nature of those treated for depression that was occurring at this time. Whereas Lewis's inpatient sample reflected the standard clinical population of depressive patients early in the twentieth century, gradually over the course of the century outpatient psychiatric clinics became the most common settings for treatment for depression. Outpatients, however, presented a far wider range of problems, including substantial numbers of normal sadness states, than the more homogeneous groups of severely disordered inpatients that Kraepelin and Lewis studied.[45] "Psychiatrists today," summarized psychiatrist Hagop Akiskal shortly before the publication of the *DSM-III* in 1980, "are faced with a large number of individuals who are seeking help for poorly defined states of psychic malaise and dysphoria that seem to defy further characterization. . . . Hence the growing vagueness of neurotic depression is paralleled by its increasing clinical visibility."[46] Extending symptom-based diagnostic methods from inpatient settings to far more heterogeneous outpatient clinics, without the simple contextual distinctions used in the past to distinguish the normal from the disordered, created the potential for unprecedented numbers of false-positive diagnoses of depressive disorder.

By the 1970s, a "hodgepodge of competing and overlapping systems" that contrasted psychotic and neurotic, endogenous and reactive, bipolar and unipolar, and many other types characterized the literature on depression.[47] Aside from a consensus that psychotic (or endogenous) depressions were distinct from neurotic states, there was virtually no agreement on the nature of nonpsychotic depressions. Researchers did not agree on whether nonpsychotic depressions were continuous or discontinuous with psychotic forms, on the one hand, or with normality, on the other. They disputed how many forms neurotic conditions took and even whether they had any distinct forms at all. Nor was it known whether some milder forms of depression were early indicators of eventual psychotic forms. In addition, little consensus existed about the particular symptoms that were essential to definitions of nonpsychotic forms of depression. Summarizing the situation in the United States and Great Britain in the mid- to late 1970s, physicians Christopher Callahan and German Berrios noted that "psychiatric diagnostic categories are at best subjective and probably irrelevant."[48] In 1980, responding to this period of confused debate characterized by the highly unsettled state of empirical findings and lack of definitive theory about the nature of non-psychotic depression, psychiatry would nonetheless adopt a definitive set of sympto
stable until the present.

Paving the Road to th
The Feighner Criteria

The proximate origins of the *1*
research psychiatrists at Washi
long as the system of classificat
was no hope for psychiatry to b
nent psychiatrists—Eli Robins ar
neo-Kraepelinian research tradi
they wanted to remedy the confi
different researchers. The St. Lou
of having agreed-on criteria tha
the basis for research studies and

In 1972, based on discussions a
diagnostic criteria that might be u
University, John Feighner, codifiec
tal disorders, including primary a
to be called the *Feighner criteria.*[49] T
lated for everyday clinical use. Ratl
of the multiplicity of different imp
possible more cumulative, compa
goal was a "common ground for di
diagnostic criteria by a number o
problem of whether patients descr
first and crucial taxonomic step sh

The Feighner criteria divided pr
depression and mania; we conside
of depression required satisfaction
dysphoric mood marked by sympt
dent, or hopeless. Second, at least fi
a total of six for definitive diagnosis
for probable diagnosis) from amon
ficulty, loss of energy, agitation, loss
slow thinking, and recurrent suicida
lasted at least 1 month and not be d

Of people who met these symp
threatening or incapacitating medic
sis of primary depressive disorder. O
was based on the fact that being inter
nesses. However, it turns out that t

adopt a definitive set of symptomatic criteria for depression that have remained stable until the present.

Paving the Road to the DSM-III: *The Feighner Criteria*

The proximate origins of the *DSM-III* criteria lie in the work of a group of research psychiatrists at Washington University in St. Louis who felt that as long as the system of classification remained without precise definitions, there was no hope for psychiatry to become a scientific discipline. Led by two prominent psychiatrists—Eli Robins and Samuel Guze—the group was inspired by the neo-Kraepelinian research tradition of analyzing symptoms statistically, and they wanted to remedy the confusing and divergent definitions of disorders by different researchers. The St. Louis group emphasized the scientific importance of having agreed-on criteria that primarily used symptomatic presentations as the basis for research studies and diagnostic decisions.

In 1972, based on discussions among the faculty regarding how to improve the diagnostic criteria that might be used in their research, a resident at Washington University, John Feighner, codified and published diagnostic criteria for 15 mental disorders, including primary and secondary affective disorders, in what came to be called the *Feighner criteria.*[49] The Feighner criteria were not explicitly formulated for everyday clinical use. Rather, they were an attempt to relieve researchers of the multiplicity of different imprecise definitions then in use and thus to make possible more cumulative, comparable, and reproducible research. The stated goal was a "common ground for different research groups. . . . The use of formal diagnostic criteria by a number of groups . . . will result in a resolution of the problem of whether patients described by different groups are comparable. This first and crucial taxonomic step should expedite psychiatric investigation."[50]

The Feighner criteria divided primary affective disorders into two categories, depression and mania; we consider only the "depression" category. Diagnosis of depression required satisfaction of three criteria. First, the patient must have dysphoric mood marked by symptoms such as being depressed, sad, despondent, or hopeless. Second, at least five additional symptoms must be present (i.e., a total of six for definitive diagnosis; four additional symptoms, or a total of five for probable diagnosis) from among a list including loss of appetite, sleep difficulty, loss of energy, agitation, loss of interest in usual activities, guilt feelings, slow thinking, and recurrent suicidal thoughts. Finally, the condition must have lasted at least 1 month and not be due to another preexisting mental disorder.

Of people who met these symptomatic criteria, only those who had life-threatening or incapacitating medical illnesses were excluded from the diagnosis of primary depressive disorder. One might have thought that this exclusion was based on the fact that being intensely sad is often a normal response to such illnesses. However, it turns out that these patients' symptoms simply warranted

a different diagnosis: that of secondary affective disorder. This category encompassed all conditions that met the same symptomatic criteria as primary disorders but that occurred with a preexisting nonaffective psychiatric illness or a life-threatening or incapacitating medical illness. Thus there were, in fact, no exclusions from disorder whatever for those who satisfied symptomatic criteria.

The Feighner criteria for affective disorders differed in significant ways from the criteria in prior empirical research on depression and, in some ways, were in tension with that research. First, all depressive conditions that did not have manic features and that were not preceded by other psychiatric or medical conditions were grouped into a single category. This system conformed to Kraepelin's theory that depression was a unitary disorder but ignored the vast majority of empirical studies that suggested possible distinctions in depressive symptomatic profiles between psychotic unipolar (i.e., not involving mania) depressions and neurotic depressions. However, we have seen that the research was not conclusive and that no consensus existed about possible distinctions between types of depressive disorder.

Where the Feighner criteria most unjustifiably deviated from considered psychiatric judgment was in making no room at all for depressive reactions of more than 1 month in duration that stemmed from normal loss responses, even including bereavement. The criteria did not allow for the possibility that some depressive symptoms were proportionate to their provoking causes even if they lasted a month, whereas others stemmed from dysfunctions. This set a crucial precedent for subsequent criteria sets that built on the Feighner work.

Why the Feighner group ignored the obvious problem of normal sadness in their criteria remains unclear. One possibility is that, to ensure that researchers would widely use the criteria, they fervently strove to avoid any inference about causation in their definitions; they might have concluded that the distinction between normality and disorder implied a particular etiological approach to classification.[51] Another is that they developed the criteria with research samples whose members clearly had some disorder and assumed the criteria would generally be used with similar samples. Or perhaps they simply were following the research tradition that immediately preceded them, which relied on statistical analysis of symptoms without regard to context.

A further possibility is that the Feighner group implicitly recognized the disordered—nondisordered-sadness distinction but assumed that intense sadness of more than 1 month's duration is "prolonged" in Hippocrates' sense and, if it involves the specified number of symptoms, is inherently disproportionate to any possible stressor and thus almost certainly disordered. If so, previous clinical observers did not accept this assumption, and it seems to conflict with the trajectory of normal response to major losses documented in chapter 2; even the *DSM-III* was to allow 2 months of normal symptomatic response to loss of a loved one. In any event, we will see that the *DSM-III* lowered this duration threshold to the much less plausible criterion of 2 weeks. In sum, unlike Kraepelin, Feighner and colleagues provide no backg
disorder from nondisorder, nor do
satisfying depressive symptom crit

How did the Feighner group de
One of the ironies of psychiatric l
the Feighner criteria's symptom-b
DSM-III was their claimed ground
retical speculation.[52] Yet, judging b
for depressive disorder, at least, had
The article references only four pu
criteria. (A fifth reference cites an u
workshop at the National Institute
six citations to publications on mar

One referenced article asserts t
condition of involutional (i.e., pos
tomatically distinguishable from
which Kraepelin had vacillated), a
eral adequacy of symptomatic cri
tients into clinical entities by symp
never been clear where the dividin
psychiatry."[53] Two other references
dicated that there was some tentat
represented the core of depressive s
depression were most likely to be "p
atric disorders other than depressic
drome."[54] The findings from these
the Feighner criteria's lumping of
final reference explicitly rejected th
of depression that do not embody cc
and their normal versus pathologica

> In classifying depressive states th
> normal and pathological reaction
> eral are normal reactions to the lo
> person, money, the depressed in
> his health—and it is not always p
> reactions from pathological dep
> alone. A depression is judged to b
> cific cause for it in the patient's in
> symptoms are too severe.[55]

None of the citations that the Feig
orts the assumption that purely syn

and colleagues provide no background understanding of how to distinguish disorder from nondisorder, nor do they state the need to evaluate whether those satisfying depressive symptom criteria are indeed disordered.

How did the Feighner group develop their influential criteria for depression? One of the ironies of psychiatric history is that the later justification for using the Feighner criteria's symptom-based diagnostic categories as the model for the *DSM-III* was their claimed grounding in empirical research rather than in theoretical speculation.[52] Yet, judging by the citations the article provides, the criteria for depressive disorder, at least, had little empirical support in the prior literature. The article references only four published articles as sources for the depression criteria. (A fifth reference cites an unpublished paper by Robins and Guze from a workshop at the National Institute of Mental Health. In addition, the work gives six citations to publications on mania, which we do not consider here.)

One referenced article asserts that there is no evidence that the particular condition of involutional (i.e., postmenopausal) depressive syndrome is symptomatically distinguishable from other depressive disorders (a question on which Kraepelin had vacillated), and it concludes with a challenge to the general adequacy of symptomatic criteria: "Attempting to group psychiatric patients into clinical entities by symptom pictures has been frustrating as it has never been clear where the dividing lines belong. This is a serious problem in psychiatry."[53] Two other references, coming out of a single research project, indicated that there was some tentative evidence for an endogenous factor that represented the core of depressive symptoms but that the symptoms of reactive depression were most likely to be "phenomenological manifestations of psychiatric disorders other than depression which 'contaminate' the depression syndrome."[54] The findings from these studies, if anything, actually contradicted the Feighner criteria's lumping of endogenous and reactive conditions. The final reference explicitly rejected the use of purely symptom-based definitions of depression that do not embody considerations about the causes of symptoms and their normal versus pathological status:

> In classifying depressive states the first distinction to be made is between normal and pathological reactions. Mourning and grief reactions in general are normal reactions to the loss of a love object—this may be another person, money, the depressed individual's prestige, his cherished hopes, his health—and it is not always possible to distinguish such normal grief reactions from pathological depression on phenomenological grounds alone. A depression is judged to be pathological if there is insufficient specific cause for it in the patient's immediate past, if it lasts too long, or if its symptoms are too severe.[55]

None of the citations that the Feighner article references for depression supports the assumption that purely symptom-based criteria can define depressive

disorders. These sources neither justify nor even address the validity of the specific definition of affective disorders in the criteria.

Soon after the publication of the Feighner criteria, the Washington University psychiatrists Robert Woodruff, Donald Goodwin, and Samuel Guze expanded their discussion of their new diagnostic criteria and their general approach to diagnosis in the first symptom-based psychiatric textbook, *Psychiatric Diagnosis.*[56] The chapter on diagnosis of affective disorders emphasized the importance of observing and measuring symptoms without any etiological inferences because of the poor state of knowledge about the causes of depression. This principle perhaps partly explains why the Feighner criteria did not allow even bereavement to be excluded from a diagnosis of depressive disorder.

In a section on differential diagnosis in the affective disorders, the text notes the following regarding bereavement (there is no discussion of other stressors):

> Making the distinction between grief and primary affective disorder can be difficult. However, grief usually does not last as long as an episode of primary affective disorder. . . . The majority of bereaved persons experience fewer symptoms than do patients with primary affective disorder. Furthermore, some symptoms common in primary affective disorder are relatively rare among persons experiencing bereavement, notably fear of losing one's mind and thoughts of self-harm.[57]

Supporting their points about the differences between bereavement and depressive disorder, Woodruff and colleagues cite several articles by psychiatrist Paula Clayton and her colleagues that document the type and duration of depressive symptoms occurring in bereavement.[58] In fact, Clayton found that after 1 month, which was the Feighner duration threshold for diagnosis of a disorder, about 40% of bereaved individuals display full *DSM*-level symptoms. Yet there is little plausibility and no scientific evidence that such a large percentage of the bereaved become disordered. Given the enormous number of individuals who experience bereavement over time, the notion that a "majority" do not experience as many symptoms as the Feighner criteria require of the disordered at the 1-month mark, and the notion that "usually" bereavement at that intense level does not last as long as the Feighner's 1-month requirement, there is no greatly reassuring evidence of its validity. Indeed, it seems to leave the door open to large numbers of false positive diagnoses of the normally bereaved, an issue that goes unaddressed.

The authors' apparent assumption that a 1-month duration and five-symptom threshold for "probable" disorder (six for "definite" diagnosis) was sufficient to discriminate disorder from normal bereavement is unwarranted on the basis of the very studies that they themselves cite. In any event, the text offers no substantive new empirical support for the proposed criteria

for depressive disorder, leaving t
challengeable as before. This text
in shaping the subsequent *DSM-I*

Meanwhile, the Feighner cri
community; by 1989, the article
cited article in the history of psyc
depressive disorder set the stage f
diagnoses, despite the fact that by
guishing intense normal from disc

The Research Diagnostic

Robert Spitzer was the major trar
what were to become the clinical
search Diagnostic Criteria (RDC),
Eli Robins of the Washington Univ
bridge between these two landma
RDC, Spitzer also developed one o
depression, the Schedule for Affect
early step toward the development
be used in epidemiologic studies th
yond the clinic to community samp

At the behest of the National I
and his colleagues developed the
reliability of psychiatric diagnoses
ogy of depression diagnoses. Like th
aimed at facilitating research, but
see. Building on the Feighner sym
15 diagnoses of the Feighner crite
types of disorder.[63]

The symptom criteria for Major
episode lasting at least 2 weeks, the
phoric mood or pervasive loss of inte
symptoms (four for a probable diagr
because of the disorder, and the abse
The major changes in the RDC from
pervasive loss of interest or pleasure
a necessary condition (reflecting a g
sure is central to depression); that s
instead of 1 month (an unexplaine
tion that potentially allowed for mar
individuals but was to find its way ir
to have either sought help from som

even address the validity of the
criteria.
riteria, the Washington University
win, and Samuel Guze expanded
a and their general approach to di-
c textbook, *Psychiatric Diagnosis.*[56]
rs emphasized the importance of
any etiological inferences because
uses of depression. This principle
iteria did not allow even bereave-
ssive disorder.
the affective disorders, the text
(there is no discussion of other

primary affective disorder can
ot last as long as an episode of
ty of bereaved persons experi-
ith primary affective disorder.
1 primary affective disorder are
g bereavement, notably fear of
m.[57]

ences between bereavement and
cite several articles by psychiatrist
nent the type and duration of de-
.[58] In fact, Clayton found that after
reshold for diagnosis of a disorder,
ll *DSM*-level symptoms. Yet there
hat such a large percentage of the
mous number of individuals who
n that a "majority" do not experi-
iteria require of the disordered at
tally" bereavement at that intense
1-month requirement, there is no
eed, it seems to leave the door open
f the normally bereaved, an issue

t a 1-month duration and five-
er (six for "definite" diagnosis)
1 normal bereavement is unwar-
hey themselves cite. In any event,
support for the proposed criteria

for depressive disorder, leaving the validity of the new criteria as empirically challengeable as before. This text, now in its fifth edition, was highly influential in shaping the subsequent *DSM-III.*[59]

Meanwhile, the Feighner criteria clearly served a need in the research community; by 1989, the article in which they appeared was the single most cited article in the history of psychiatry.[60] Their widely influential definition of depressive disorder set the stage for psychiatry's use of purely symptom-based diagnoses, despite the fact that by nature this approach was incapable of distinguishing intense normal from disordered responses.

The Research Diagnostic Criteria

Robert Spitzer was the major translator of the Feighner research criteria into what were to become the clinical diagnostic criteria of the *DSM-III.* The Research Diagnostic Criteria (RDC), which Spitzer created in collaboration with Eli Robins of the Washington University group and published in 1978, was the bridge between these two landmark achievements.[61] In conjunction with the RDC, Spitzer also developed one of the first structured interviews to measure depression, the Schedule for Affective Disorders and Schizophrenia (SADS), an early step toward the development of structured questionnaires that would later be used in epidemiologic studies that applied the new diagnostic approach beyond the clinic to community samples (see chapter 6).[62]

At the behest of the National Institute of Mental Health (NIMH), Spitzer and his colleagues developed the RDC to overcome concerns about the low reliability of psychiatric diagnoses and to create a more sophisticated typology of depression diagnoses. Like the Feighner criteria, the RDC were explicitly aimed at facilitating research, but their clinical application was not hard to see. Building on the Feighner symptom-based approach, they expanded the 15 diagnoses of the Feighner criteria to 25 major types and many more subtypes of disorder.[63]

The symptom criteria for Major Depressive Disorder in the RDC required an episode lasting at least 2 weeks, the presence of a prominent and persistent dysphoric mood or pervasive loss of interest or pleasure, five out of eight additional symptoms (four for a probable diagnosis), help seeking or impaired functioning because of the disorder, and the absence of features that suggest schizophrenia. The major changes in the RDC from the Feighner criteria were stipulations that pervasive loss of interest or pleasure could be substituted for dysphoric mood as a necessary condition (reflecting a growing view that loss of capacity for pleasure is central to depression); that symptoms need only be present for 2 weeks instead of 1 month (an unexplained substantial reduction in required duration that potentially allowed for many more false positive diagnoses of normal individuals but was to find its way into the *DSM-III*); and that the patient had to have either sought help from someone or have impaired social functioning

(essentially an early form of the later clinical significance criterion). A number of exclusion criteria that eliminated those with schizophrenia from a depression diagnosis were also added, as were 11 subtypes of MDD. (The nonmutually exclusive subtypes of MDD, the original motivation for NIMH's interest, were primary, secondary, recurrent unipolar, psychotic, incapacitating, endogenous, agitated, retarded, situational, simple, and predominant mood.) Despite the lowering of both the duration and symptom thresholds from the Feighner criteria to levels that would later be incorporated into the *DSM-III*, the RDC's criteria for MDD contained no exclusions for bereavement or any other normal reaction, although they did require researchers to ascertain during their interviews with patients whether bereavement was present.[64]

For reasons we consider in the next section, a major concern in constructing the RDC was reliability of diagnosis, that is, whether different diagnosticians would come to the same diagnosis based on the same information. Studies using the RDC indicated great overall success in achieving reliability. For Major Depressive Illness, the initial reports indicated the remarkable reliability of .97.[65] Other reports indicated reliabilities of about .90.[66] Many considered the apparent improvement of reliability to be a great advance, as the remarks of the noted diagnostician Alvin Feinstein indicate:

> The production of operational identifications has been a pioneering, unique advance in nosology. . . . In the field of diagnostic nosology, the establishment of operational criteria represents a breakthrough that is as obvious, necessary, fundamental, and important as the corresponding breakthrough in obstetrics and surgery when Semmelweis, Oliver Wendell Holmes, and, later on, Lord Lister, demanded that obstetricians and surgeons wash their hands before operating on the human body.[67]

We will see that Feinstein's enthusiasm for Spitzer's accomplishments reflects what was to become Spitzer's greatest achievement, his shepherding of the creation of an entirely new psychiatric clinical diagnostic classification system using the same principles as the RDC to ensure reliability. However, we sound a preliminary caution in anticipation of the discussion of the *DSM-III:* It is true that when symptoms alone are the basis for diagnoses, people can be trained to apply the criteria according to rules and thus to agree, and the reliability of diagnoses may well increase. But are the agreed judgments correct in identifying disorders (i.e., valid)? These studies did not assess the validity of the diagnosis in predicting course, response to therapies, or etiology of depressive conditions. Moreover, the RDC and the Feighner criteria did not involve any systematic attempt to distinguish normal intense sadness from depressive disorder, casting further doubt on the validity of these approaches. Introducing judgments about normal versus disordered reactions to circumstances into diagnostic criteria is challenging to do and would likely lessen reliability, but even so it might substantially enhance validity. To this day, we shall see challenge.

The DSM-III *as a Respo Confronting Psychiatry*

The publication of the *DSM-III* i the history of psychiatric diagnc seen beforehand as particularly ir from advocates of different theor cess. Spitzer's work on the comn prominent role in brokering the r and his development of the RDC c *DSM-III* task force. Spitzer used tl nostic system that reflected previ psychiatry more scientific.[69]

The *DSM-III* revolution directl Feighner criteria and RDC into the embraced symptom-based diagnos the translation of research criteria the diagnostic criteria must reflect " search.[70] His role required not only also those of a master politician at among various clinical constituen system threatened their traditional

But what motivated Spitzer tc symptom-based definitional approa why did clinicians, who are concern interest in reliable classification syste classification system that had emerg

It turns out that the new system a cians, as well as researchers, faced influence had waned. The psychiat theoretical schools, and different clir fundamental nature, causes, and tre agnostic manual, therefore, had to be perspectives. The lists of explicit symp liability but also were *theory neutral* in particular theory of the cause of psycl The new criteria were *descriptive* rathe postulated psychodynamic causes of against anxiety). Defining disorders etiology, turned out to be a useful too

validity. To this day, we shall see, psychiatry has not adequately addressed this challenge.

The DSM-III *as a Response to the Challenges Confronting Psychiatry*

The publication of the *DSM-III* in 1980 is justifiably viewed as a watershed in the history of psychiatric diagnosis.[68] Yet the revision of the *DSM-II* was not seen beforehand as particularly important, and there was no political jockeying from advocates of different theoretical perspectives to be in control of the process. Spitzer's work on the committee charged with revising the *DSM-II*, his prominent role in brokering the removal of homosexuality from that manual, and his development of the RDC criteria led to his appointment as chair of the *DSM-III* task force. Spitzer used the opportunity to create a new kind of diagnostic system that reflected previous decades of thought about how to make psychiatry more scientific.[69]

The *DSM-III* revolution directly incorporated many of the features of the Feighner criteria and RDC into the official psychiatric nosology and specifically embraced symptom-based diagnostic criteria. Spitzer himself recognized that the translation of research criteria into a manual for clinical use required that the diagnostic criteria must reflect "clinical wisdom" as well as evidence from research.[70] His role required not only the skills of a knowledgeable researcher but also those of a master politician attempting to mollify and to find compromise among various clinical constituencies that felt that the new symptom-based system threatened their traditional diagnostic practices.

But what motivated Spitzer to borrow so heavily from the RDC-style symptom-based definitional approach to diagnosis in revising the *DSM?* And why did clinicians, who are concerned with treating individuals and have little interest in reliable classification systems for research, accept the symptom-based classification system that had emerged from the Feighner criteria and RDC?

It turns out that the new system addressed several major problems that clinicians, as well as researchers, faced at the time. By the 1970s, psychoanalytic influence had waned. The psychiatric profession was divided into numerous theoretical schools, and different clinicians shared few assumptions about the fundamental nature, causes, and treatments of mental disorders. The new diagnostic manual, therefore, had to be serviceable for clinicians of many varying perspectives. The lists of explicit symptoms in the *DSM-III* not only improved reliability but also were *theory neutral* in the sense that they did not presuppose any particular theory of the cause of psychopathology, psychoanalytic or otherwise. The new criteria were *descriptive* rather than *etiological* and purged references to postulated psychodynamic causes of a disorder (e.g., internal conflict, defense against anxiety). Defining disorders on the basis of symptoms, regardless of etiology, turned out to be a useful tool in gaining the acceptance of clinicians

of varying allegiances who could at least feel that all factions were on a level playing field in using the theory-neutral definitions.

Moreover, psychiatric diagnoses were under attack from a variety of sources. Behaviorists claimed that all behavior, including psychopathology, is the result of normal learning processes and thus that no mental disorders in the medical sense really exist.[71] The "antipsychiatry" movement, inspired by writers such as psychiatrist Thomas Szasz and sociologist Thomas Scheff, portrayed psychiatric diagnosis as a matter of using medical terminology to apply social control to undesirable but not truly medically disordered behavior.[72]

In addition, by 1980 private and public third parties were financing most medical treatment.[73] The murky unconscious entities of the *DSM-II* and the erosion of psychiatry's medical legitimacy did not provide a solid basis for insurance reimbursement. Although no evidence indicates that insurers influenced the development of the symptom-based disorders of the manual, the new diagnoses provided a better fit with the goal of third parties to reimburse the treatment of only specific diseases. On reflection, clinicians may not have agreed with some features in the new manual, such as the abandonment of contextual criteria, but they realized that the new system had many benefits for them.

Most pressing of all was an erosion of the credibility of psychiatry due to attacks on the meaningfulness of diagnosis. Although he had psychoanalytic training, Spitzer, like the St. Louis group, saw unverified theory and resistance to empirical testing as the major obstacles to psychiatry's attaining scientific status.[74] The central element in Spitzer's vision of psychiatry, pursued in his prodigious research efforts in the 1960s and 1970s and culminating in the *DSM-III* in 1980, was the development of a *reliable* system of classification in which different diagnosticians would generally arrive at the same diagnosis based on the same clinical information.[75]

Because the *DSM-II* did not provide specific symptoms that determined psychiatric diagnoses, psychiatrists were forced to use their own clinical judgments in assessing how well each patient fit a particular diagnosis. This led to great disparities in the application of diagnostic labels. For example, the well-known U.S.-U.K. Diagnostic Project, the results of which were published in 1972, studied the ways that psychiatrists in these two countries diagnosed mental disorders. The study demonstrated an alarming lack of agreement between American and British psychiatrists and among psychiatrists within each group. For example, more than five times as many British as American psychiatrists made diagnoses of depressive disorders.[76]

In addition to the U.S.-U.K. study, a great number of studies generally showed remarkable lack of diagnostic agreement in cases in which psychiatrists received the same information (e.g., a videotaped clinical interview).[77] These studies challenged the reliability not only of distinguishing closely related diagnostic categories, such as one affective disorder from another, but also of distinguishing between larger categories, such as affective versus anxiety disorders, and

between overall types of disorde
psychosis versus normality.

Perhaps the most dramatic ar
mark in the critique of psychiatr
psychiatrists to distinguish norn
David Rosenhan published a stu
eight normal individuals present
auditory hallucinatory symptom
like "thud," "dull," and "empty'
All of these pseudo-patients were
all as schizophrenic), and they ren
even though they immediately rev
however, did identify several pseud

To get the flavor of the views p
tences in the introduction to Roser

> Normality and abnormality, sa
> flow from them may be less subs
> Based in part on theoretical and
> on philosophical, legal, and the
> psychological categorization of r
> right harmful, misleading, and p
> in this view, are in the minds of
> characteristics displayed by the c

Based on his results, Rosenhan c
guish the sane from the insane in ps

The threat of such gross invalidity
Rosenhan's participants would, unde
mal) was not only an acute embarrass
to the scientific status of psychiatry.
the methodological flaws in Rosenha
show only that Rosenhan had not pr
by its nature flawed; it could not dem
had an adequately reliable diagnostic
was to be devoted to the project of crea

Although acknowledging that a
Spitzer emphasized that validity requ
would categorize different syndrome
course and response to treatment.[80]
even agree on the diagnosis, then cle
accurate, and there must be low over
reliability of diagnoses across setting

between overall types of disorder, such as psychosis versus neurosis or even psychosis versus normality.

Perhaps the most dramatic and influential such study, now seen as a landmark in the critique of psychiatric diagnosis, directly challenged the ability of psychiatrists to distinguish normality from psychosis. In 1973, psychologist David Rosenhan published a study in the prestigious journal *Science* in which eight normal individuals presented themselves at hospitals and reported only auditory hallucinatory symptoms (they claimed to hear a voice saying things like "thud," "dull," and "empty"), otherwise acting and speaking normally. All of these pseudo-patients were admitted and classified as psychotic (almost all as schizophrenic), and they remained so classified for various periods of time, even though they immediately reverted to normal behavior. Hospital residents, however, did identify several pseudo-patients as likely normals.

To get the flavor of the views prominent at the time, consider a few of sentences in the introduction to Rosenhan's article:

> Normality and abnormality, sanity and insanity, and the diagnoses that flow from them may be less substantive than many believe them to be. . . . Based in part on theoretical and anthropological considerations, but also on philosophical, legal, and therapeutic ones, the view has grown that psychological categorization of mental illness is useless at best and downright harmful, misleading, and pejorative at worst. Psychiatric diagnoses, in this view, are in the minds of observers and are not valid summaries of characteristics displayed by the observed.[78]

Based on his results, Rosenhan concluded: "It is clear that we cannot distinguish the sane from the insane in psychiatric hospitals."

The threat of such gross invalidity and, by implication, unreliability (for surely Rosenhan's participants would, under other circumstances, have been judged normal) was not only an acute embarrassment to clinical expertise but also a challenge to the scientific status of psychiatry. Spitzer himself wrote a scathing critique of the methodological flaws in Rosenhan's study.[79] However, such a critique could show only that Rosenhan had not proved his claim that psychiatric diagnosis is by its nature flawed; it could not demonstrate that psychiatric diagnosis, in fact, had an adequately reliable diagnostic system. Much of Spitzer's subsequent effort was to be devoted to the project of creating and nurturing such a system.

Although acknowledging that a reliable system is not necessarily valid, Spitzer emphasized that validity requires reliability. A valid diagnostic system would categorize different syndromes accurately and thereby ought to predict course and response to treatment.[80] But if different diagnosticians could not even agree on the diagnosis, then clearly many of their diagnoses must be inaccurate, and there must be low overall diagnostic validity. Moreover, without reliability of diagnoses across settings, cumulative research could not proceed

effectively. Therefore, the primary goal of the psychiatric profession had to be the development of a clear system of diagnostic rules that specified inclusion and exclusion criteria for each diagnosis and promoted a high degree of inter-judge agreement. Even if lacking in validity, such a reliable system could provide a scientifically adequate starting place from which researchers could bootstrap themselves to a more valid system.

However, as many concerned critics pointed out, just creating a reliable system that has clear rules that everybody can follow does not ensure even an approximation of validity; unless the rules are accurate, the reliability might just represent everybody together getting the same wrong answer![81] For example, if symptoms of intense sadness are used to indicate depressive disorder, such symptoms might be identified reliably, but the vast majority of conditions so recognized might not, in fact, be disorders. The field trials conducted before the publication of the *DSM-III*, in which hundreds of psychiatrists had tested the empirical adequacy of the diagnoses, did not compare the effectiveness of symptom-based criteria sets with other alternative ways of conceptualizing depression.[82] They tested only whether different psychiatrists could use the criteria in the same way but did not establish whether they were valid indicators of disorder. As one of Spitzer's collaborators notes: "pathologic conditions (were) redefined *before* empiric investigation (was) conducted."[83] And it is far from certain that such a system, if seriously invalid to begin with, would automatically evolve into a valid system. The implication is that considerations of validity cannot be entirely placed on the back burner while issues of reliability are resolved; both must be pursued together, and the two must inform each other in order to approach more reliable judgments that are also valid.

Between psychiatry's theoretical fragmentation, its diagnostic unreliability, and the antipsychiatry critique, not only psychiatry's claim to scientific status but even its legitimacy as a medical field seemed in jeopardy. The specific criteria of the *DSM-III* appeared to meet these challenges and place the field on a more sound scientific footing. In one fell swoop, Spitzer's incorporation of symptom-based operational definitions of disorders into the *DSM* managed to confront a range of challenges to psychiatry and to facilitate an about-face in psychiatry's status and fortunes, especially coinciding as it did with the advent of new medications that were also bolstering the status of the psychiatric profession.

But even a justified revolution has some unwarranted casualties. Having considered the nature and reasons for the *DSM-III* revolution in general, we now turn to the *DSM-III* criteria for depressive disorder.

The *DSM-III*'s Approach to Depressive Disorder

The *DSM-III* criteria for depression almost completely mirrored the approaches of the Feighner and RDC criteria (the next chapter discusses in detail the similar

DSM-IV criteria). They used sym
doned or demoted etiological co
as neurotic versus psychotic and
ferent categories of diagnosis. L
earlier *DSMs*, the *DSM-III* rejecte
insanity and depression, instead
"major depression," from "bipola
area of controversy, family studi
of medication responses had all
grand unification of affective dis
though MDD covered psychotic c
tions comprised only a small min
depression" had come to be the p
the manual.

Likewise, the *DSM-III* abandor
versus proportionate reactions to
object or cherished possession."
egories of disorder in the *DSM-I*
fiers such as "excessive" or "unr
responses. Yet the *DSM-III* distin
of symptoms regardless of their
exception of the bereavement exc

The logic behind the bereav
improvement over the Feighner
otherwise meet symptomatic crit
normal and transient responses
from the work of Paula Clayton,
versity group and of the *DSM-II*
had shown that depressive-like s
bereavement but that they usua
noted earlier, Woodruff, Goodwir
not incorporate it into their diag
incorporate Clayton's findings ir
did not apply the exclusion to the
have the same features as berea
tion, ill health, or financial reve
other stressors simply never came
orders work group as a possible l
sions seems to have been a by-pro
exclusionless Feighner and RDC c
spirit of the *DSM-III* effort.

Various reasons have been cit
sions from major depression for n

DSM-IV criteria). They used symptoms to specify depressive disorder and abandoned or demoted etiological concepts, as well as traditional distinctions such as neurotic versus psychotic and endogenous versus reactive as a basis for different categories of diagnosis. Like the Feighner criteria and RDC, as well as earlier *DSM*s, the *DSM-III* rejected Kraepelin's unification of manic-depressive insanity and depression, instead distinguishing unipolar depressive disorder, or "major depression," from "bipolar" disorders. Although this remains an active area of controversy, family studies, clinical observations, and distinct patterns of medication responses had all served to thoroughly undermine Kraepelin's grand unification of affective disorders long before the *DSM-III*. Moreover, although MDD covered psychotic depression, it was understood that such conditions comprised only a small minority of those falling under the criteria; "simple depression" had come to be the predominant form of depression of concern in the manual.

Likewise, the *DSM-III* abandoned the *DSM-II* distinction between "excessive" versus proportionate reactions to an "identifiable event such as the loss of a love object or cherished possession." This is surprising given that many other categories of disorder in the *DSM-III*, such as some anxiety disorders, use qualifiers such as "excessive" or "unreasonable" to separate disorders from normal responses. Yet the *DSM-III* distinguishes depressive disorders solely on the basis of symptoms regardless of their relationship to circumstances, with the single exception of the bereavement exclusion.

The logic behind the bereavement exclusion, which represents a major improvement over the Feighner criteria and RDC, is that states of grief that otherwise meet symptomatic criteria are not disorders because they represent normal and transient responses to loss. The exclusion seems to have resulted from the work of Paula Clayton, a prominent member of the Washington University group and of the *DSM-III* Task Force on Affective Disorders. Her work had shown that depressive-like symptoms commonly arose during periods of bereavement but that they usually remitted after a fairly short time.[84] As we noted earlier, Woodruff, Goodwin, and Guze mentioned Clayton's work but did not incorporate it into their diagnostic criteria for depression. The *DSM-III* did incorporate Clayton's findings in developing the bereavement exclusion, but did not apply the exclusion to the reactions to any other types of loss that may have the same features as bereavement, such as reactions to marital dissolution, ill health, or financial reversal. So far as we can ascertain, reactions to other stressors simply never came up for discussion by the *DSM-III* affective disorders work group as a possible basis for exclusions.[85] The lack of such exclusions seems to have been a by-product of deriving the *DSM-III* criteria from the exclusionless Feighner and RDC criteria and the symptom-oriented diagnostic spirit of the *DSM-III* effort.

Various reasons have been cited to justify the *DSM*'s failure to allow exclusions from major depression for normal situations other than bereavement. For

one thing, such exclusions could pose a serious challenge to reliability; other stressors often lack the relatively clear-cut nature of bereavement, and it would be more difficult to measure their magnitude and to judge their proportionality to the resulting response. However, as we have noted, it makes no sense for reliability to trump validity in constructing diagnostic criteria. In any event, the framers of the *DSM-III*, in creating criteria for "complicated" bereavement, discussed in the next chapter, showed that it is possible to reflect such subtle distinctions within a given stressor type. Similar efforts could have been made to provide guidelines for when reactions to other major stressors represent normal versus disordered reactions.

The question of whether sadness is a proportionate response to real loss is sometimes argued to be an etiological issue that has no place in a theory-neutral manual.[86] But this objection is based on confusion about the nature and point of theory neutrality. The distinction between normal, proportional responses to events and disorders in which sadness derives from an internal dysfunction is not really a theory-laden distinction in the sense relevant to the *DSM-III*'s need for theory neutrality. Different theories offer different accounts—whether biological, psychodynamic, behavioral, cognitive, or social—of the nature and etiology of the dysfunction that underlies depressive disorder, and a theory-neutral manual must not accept one theory over another as part of the definition of the disorder. It can, however, acknowledge that all etiological theories share the notion of normal, proportional responses versus dysfunction-based responses. After all, medical thinkers from Aristotle to Kraepelin understood this notion in more or less the same way, and it identifies the common target that rival theories attempt to explain. This distinction is not an etiological hypothesis of the kind that a theory-neutral manual needs to exclude.

Another objection to considering the broader contexts of depressive responses in the *DSM-III* might have stemmed from the impression that psychotropic medication worked on all unipolar depressions, irrespective of the relation to triggering events, so that the "with cause" versus "without cause" distinction was irrelevant to treatment decisions, at least among hospitalized depressives.[87] However, even if medication sometimes works with normal reactions, the normality-versus-disorder distinction can have important prognostic implications for how aggressively to treat a condition and for deciding what kinds of treatments or changes in circumstances might help. Analogously, the fact that, say, Ritalin works on normal and disordered individuals alike to make them more focused, or that growth hormone makes both normal and disordered short children taller, does not imply that diagnosis can justifiably ignore the distinction between normality and disorder.

Finally, the *DSM-III*'s ignoring of normal states of intense sadness might have reflected a fear of misdiagnosing the truly disordered as normal, especially given that depressed patients are subject to suicide risk. Yet no effort was made to balance the risks of false negatives with the costs of false positives that arise from

labeling normal people as disorde
apprehension about the possible n
and of other treatments for norm
risk in some populations.[88] Majo
felt that it was important to iden
them from depressive disorders, fo
essarily ignoring a distinction, it is
exercise caution so as to err on the

Conclusion

The *DSM-III*'s largely decontextu
efforts to enhance reliability, to c
with a variety of theoretical persu
of the profession. But in the urge
part inadvertently rejected the pre
tion that explored the context an
someone is suffering from intense
unwitting result of this effort, esp
conditions of inpatients to the far
and community members, was to
ness that, ironically, can be argue
than more scientifically valid.

labeling normal people as disordered, a cost that is clearer today with the growing apprehension about the possible negative side effects of antidepressive medication and of other treatments for normal sadness, including potential increased suicide risk in some populations.[88] Major psychiatric theoreticians prior to the *DSM-III* felt that it was important to identify normal cases of sadness and to distinguish them from depressive disorders, for good reason. Rather than entirely and unnecessarily ignoring a distinction, it is more prudent to simply use it when helpful but exercise caution so as to err on the side of safety in applying the distinction.

Conclusion

The *DSM-III*'s largely decontextualized, symptom-based criteria stemmed from efforts to enhance reliability, to develop a common language for psychiatrists with a variety of theoretical persuasions, and to bolster the scientific credentials of the profession. But in the urgent quest for reliability, the criteria for the most part inadvertently rejected the previous 2,500 years of clinical diagnostic tradition that explored the context and meaning of symptoms in deciding whether someone is suffering from intense normal sadness or a depressive disorder. The unwitting result of this effort, especially as psychiatry turned from the serious conditions of inpatients to the far more heterogeneous conditions of outpatients and community members, was to be a massive pathologization of normal sadness that, ironically, can be argued to have made depressive diagnosis less rather than more scientifically valid.

5
Depression in the *DSM-IV*

We claimed in chapter 1 that a flawed definition may be facilitating the recent surge in reported depressive disorder and may even lie at its very heart. To justify our claim, we now turn to a detailed examination of the *DSM* criteria for depressive and related disorders. Although the history of depression presented in the preceding chapter logically takes us up to the *DSM-III*, in order to ensure that our discussion applies to current diagnostic practices, we address the criteria presented in the latest edition—the fourth, text-revised edition *DSM-IV-TR* (2000). This does not represent much of a conceptual leap because the current criteria are almost identical to those in the *DSM-III*.

DSM-IV Affective Disorders

We start by placing the *DSM* criteria for Major Depressive Disorder (MDD) in the context of the *DSM*'s approach to affective disorders, also known as mood disorders, the larger category under which depressive disorders fall. The following distinctions are useful to keep in mind:

Unipolar Versus Bipolar Mood Disorders

MDD is "unipolar" depression, which means that the individual has only depressive symptoms rather than oscillating back and forth between depressive and manic symptoms such as elevated mood and grandiosity. Mood disorders that include manic episodes are known as *Bipolar Disorders* (formerly *manic-depressive disorders*), which are relatively rare compared with the claimed rates of unipolar depressive disorder. Bipolar I Disorder is often quite severe; milder forms include Bipolar II Disorder and Cyclothymic Personality Disorder. None of these forms of bipolar disorder is the focus here.

Major Depressive Diso

MDD generally occurs over tim
intense episodes separated by in
toms. Another, less common fo
occurs more or less continuousl
and which is discussed later in th

Major Depressive Disor Depressive Episode

The *DSM* defines various subtype
on the pattern of occurrences of
plus some additional criteria. In
the diagnostic "action" occurs; th
informative:

Criteria for Major De

A. Presence of a Major Depre
B. The Major Depressive Epis
affective Disorder and is nc
phreniform Disorder, Delu
Otherwise Specified.
C. There has never been a Ma
manic Episode.[1]

In other words, the criteria for
ence at least one Major Depressiv
some other psychotic disorder (not
depression as long as they cannot
psychotic disorder) and is not part
manic elements. However, almost
and are not part of some other dis
criteria for MDD essentially come d
the much more informative definit

DSM-IV *Criteria for Majo*

A. Five (or more) of the follow
the same 2-week period and
tioning; at least one of the s
(2) loss of interest or pleasur

Major Depressive Disorder Versus Dysthymia

MDD generally occurs over time in a series of quasi-discrete symptomatically intense episodes separated by intervals without symptoms or with fewer symptoms. Another, less common form of depressive disorder is Dysthymia, which occurs more or less continuously for long periods of time at a less intense level and which is discussed later in this chapter.

Major Depressive Disorder Versus Major Depressive Episode

The *DSM* defines various subtypes of MDD (e.g., single episode, recurrent) based on the pattern of occurrences of what it calls Major Depressive Episodes (MDE) plus some additional criteria. In fact, it is in the criteria for MDE that most of the diagnostic "action" occurs; the criteria for MDD itself are brief and not very informative:

Criteria for Major Depressive Disorder

A. Presence of a Major Depressive Episode.
B. The Major Depressive Episode is not better accounted for by Schizoaffective Disorder and is not superimposed on Schizophrenia, Schizophreniform Disorder, Delusional Disorder, or Psychotic Disorder Not Otherwise Specified.
C. There has never been a Manic Episode, a Mixed Episode, or a Hypomanic Episode.[1]

In other words, the criteria for MDD simply require that the patient experience at least one Major Depressive Episode, and that the episode is not part of some other psychotic disorder (note that psychotic symptoms can be part of the depression as long as they cannot be better explained as indicating some other psychotic disorder) and is not part of another kind of mood disorder containing manic elements. However, almost all depressive episodes are indicative of MDD and are not part of some other disorder. Thus, in the vast majority of cases, the criteria for MDD essentially come down to the criteria for MDE. We thus examine the much more informative definition of MDE at some length.

DSM-IV *Criteria for Major Depressive Episode*

A. Five (or more) of the following symptoms have been present during the same 2-week period and represent a change from previous functioning; at least one of the symptoms is either (1) depressed mood or (2) loss of interest or pleasure.

1. Depressed mood most of the day, nearly every day, as indicated by either subjective report (e.g., feels sad or empty) or observation made by others (e.g. appears tearful). Note: In children and adolescents, can be irritable mood.
2. Markedly diminished interest or pleasure in all, or almost all, activities most of the day, nearly every day (as indicated by either subjective account or observation made by others).
3. Significant weight loss when not dieting or weight gain (e.g., a change of more than 5% of body weight in a month), or decrease or increase in appetite nearly every day. Note: In children, consider failure to make expected weight gain.
4. Insomnia or hypersomnia nearly every day.
5. Psychomotor agitation or retardation nearly every day (observable by others, not merely subjective feelings of restlessness or being slowed down).
6. Fatigue or loss of energy nearly every day.
7. Feelings of worthlessness or excessive or inappropriate guilt (which may be delusional) nearly every day (not merely self-reproach or guilt about being sick).
8. Diminished ability to think or concentrate, or indecisiveness, nearly every day (either by subjective account or as observed by others).
9. Recurrent thoughts of death (not just fear of dying), recurrent suicidal ideation without a specific plan, or a suicide attempt or a specific plan for committing suicide.

B. The symptoms do not meet criteria for a Mixed Episode.
C. The symptoms cause clinically significant distress or impairment in social, occupational, or other important areas of functioning.
D. The symptoms are not due to the direct physiological effects of a substance (e.g., a drug of abuse, a medication) or a general medical condition (e.g., hypothyroidism).
E. The symptoms are not better accounted for by Bereavement, i.e., after the loss of a loved one, the symptoms persist for longer than 2 months or are characterized by marked functional impairment, morbid preoccupation with worthlessness, suicidal ideation, psychotic symptoms, or psychomotor retardation.[2]

Anyone reporting at least five of the nine symptoms in criterion A, including at least one of depressed mood or loss of interest or pleasure, for a 2-week period is considered to have Major Depressive Episode and thus, generally, MDD. Note that even for those satisfying the symptom criteria, there are the four exclusions in Criteria B through E, eliminating the following from diagnosis: (1) conditions that also include manic symptoms, which are classified under bipolar disorders; (2) conditions that do not cause clinically significant role impairment or distress; (3) conditions that are
or use of either an illegal subst
nosed as Mood Disorder Due to G
Mood Disorder; or (4) condition
has lasted longer than 2 month
toms; this is considered a case of

How the *DSM* Criter Address the Distinct and Normal Sadnes

Symptom and Duratio

The *DSM-IV* tries to exclude norm
ders via the various features of its
old for the diagnosis sets a higher t
would meet; (2) the specific natur
herently suggest pathology, as in f
tion, or recurrent thoughts of dea
which five symptoms must cluster
individual symptoms experienced
severity, intensity, and frequency o
imal duration—for example, that
2-week period, be "marked" or "sig
a percentage weight loss—also elin

There is no question that these
many episodes of normal sadness
However, such strategies for disti
have two disadvantages. First, inc
nosis in order to eliminate false p
negatives, by which genuine disorc
a condition is not a matter of the
with a limited number of symptom

Second, although the occurren
ally more harmful, it is not alway
symptoms, or more prolonged sy
chapter 2 documented, unusually
many intense symptoms in other
symptoms that occur during norm
the depressive symptoms listed in t
disorders. Moreover, some people a
more severe normal responses to st

distress; (3) conditions that are the direct result of a general medical condition or use of either an illegal substance or prescribed medication; these are diagnosed as Mood Disorder Due to General Medical Condition or Substance-Induced Mood Disorder; or (4) conditions that stem from bereavement, unless the grief has lasted longer than 2 months or involves certain particularly severe symptoms; this is considered a case of "complicated bereavement."

How the *DSM* Criteria for Major Depression Address the Distinction Between Disorder and Normal Sadness

Symptom and Duration Criteria

The *DSM-IV* tries to exclude normal depressive conditions from diagnosis as disorders via the various features of its symptom criteria: (1) its five-symptom threshold for the diagnosis sets a higher threshold than many normal periods of sadness would meet; (2) the specific nature of some of the individual symptoms might inherently suggest pathology, as in feelings of worthlessness, psychomotor retardation, or recurrent thoughts of death; (3) the required duration of 2 weeks, during which five symptoms must cluster together, eliminates shorter periods or sporadic individual symptoms experienced discontinuously over time; and (4) the required severity, intensity, and frequency of the symptoms during at least the 2-week minimal duration—for example, that they must occur "nearly every day" during a 2-week period, be "marked" or "significant," or feature other benchmarks such as a percentage weight loss—also eliminates many milder forms of normal sadness.

There is no question that these features of the symptom criteria do eliminate many episodes of normal sadness from being mistakenly classified as disorders. However, such strategies for distinguishing disordered from normal responses have two disadvantages. First, increases in the symptomatic threshold for diagnosis in order to eliminate false positives can often inadvertently increase false negatives, by which genuine disorders go unrecognized. The disordered status of a condition is not a matter of the number of symptoms because mild disorders with a limited number of symptoms can exist.

Second, although the occurrence of a greater number of symptoms is generally more harmful, it is not always the case that more symptoms, more severe symptoms, or more prolonged symptoms imply dysfunction and disorder. As chapter 2 documented, unusually harsh environmental stressors often produce many intense symptoms in otherwise normal individuals, and the depressive symptoms that occur during normal periods of sadness are generally similar to the depressive symptoms listed in the *DSM* criteria that occur during depressive disorders. Moreover, some people are temperamentally more sensitive and have more severe normal responses to stress than others.

Thus setting high symptom thresholds in terms of number, intensity, or continuity over a 2-week period does not effectively address the dysfunction problem—that is, the problem of distinguishing whether the symptoms are part of a normal sadness reaction or are the result of a dysfunction of sadness-generating mechanisms. Intense normal sadness in response to a variety of major losses can easily include the five symptoms the *DSM* requires, such as low mood, lack of pleasure in usual activities, sleeplessness, lack of appetite, and difficulty concentrating on usual tasks. Nor is the required severity of the *DSM* symptoms, specified in some cases by qualifiers such as "recurrent," "marked," or "diminished," generally of such a distinctive level that it would characterize disordered rather than intense normal sadness responses. Likewise, the 2-week duration does not adequately distinguish potentially normal-range intense reactions to serious losses, such as the end of a marriage or a potentially terminal medical diagnosis, from depressive disorders. Normal reactions to major losses can easily last more than 2 weeks. Certainly, the severity of the symptoms themselves, having five of them, and experiencing them almost every day during a 2-week period does offer a stark contrast to usual functioning and thus may seem on first glance to impart validity. But when the contrast is between depressive disorder and periods of intense normal sadness in response to major losses, normal sadness can easily meet these requirements.

Moreover, many of the symptoms, such as difficulty sleeping and fatigue, have very high base rates in the general population in response to a variety of stresses and are not at all distinctive of depression, normal or disordered, or even of disorder in general. Thus individuals without a depressive disorder might accidentally reach the threshold due to the presence of unrelated symptoms during a period of normal low mood.

It is true that some symptoms, such as complete immobilization, a morbid and unjustified preoccupation with one's worthlessness, hallucinations, and delusions, do not significantly overlap with normal functioning. These symptoms might generally indicate dysfunctions rather than designed sadness, especially if persistent. However, the diagnosis of MDD does not require the presence of such especially severe symptoms.

Exclusion for Bereavement

One way in which the *DSM* attempts to make up for any weaknesses in the symptom criteria's ability to distinguish disorder from nondisorder is through the exclusion clauses. This is the main purpose of the bereavement exclusion. However, like every other mental or physical function, grief can "go wrong" and become disordered. For this reason, the bereavement exclusion has its own exclusion-to-the-exclusion that allows depressive symptoms associated with grief sometimes to be classified as true disorders after all. This occurs when grief responses persist for longer than 2 months, cause marked functional impairment, or include

especially severe symptoms, su
ness, suicidal ideation, psychor
is also worth noting that during
loved one's presence are not unc
pathological.)

One might dispute the 2-m
might argue that normal bereav
plicated" symptoms that the DS
far the major flaw in this exclus
normal sadness responses to an
would have been easy to genera
accompanying exclusion-to-the-
this opportunity was foregone, t
constructive attempt to validly c
is too limited to adequately addr
criteria.

*Exclusion for General M
Substance-Use-Induced*

The exclusion from MDD diagno
from the physiological effects of
shifts such cases into alternative
General Medical Condition or Su
ries, although not our focus here
For example, such disorders are
sponses to having a medical con
lems that result from using or bei
of the complex challenges practi
cate depression from similar sym
result of different disorders.

The Clinical-Significanc

Perhaps the most important attem
disordered from normal sadness r
which requires that "the symptom
ment in social, occupational, or oth
implicitly acknowledges that even
symptom criteria might still not in
dress the basic validity problems o
the negative consequences of a co
condition is to be clinically relevan

especially severe symptoms, such as morbid preoccupation with worthlessness, suicidal ideation, psychomotor retardation, or psychotic symptoms.[3] (It is also worth noting that during bereavement transient hallucinations of a lost loved one's presence are not uncommon, and they are not generally considered pathological.)

One might dispute the 2-month limit on normal bereavement, and one might argue that normal bereavement may sometimes include one of the "complicated" symptoms that the *DSM* says are sufficient for disorder. However, by far the major flaw in this exclusion criterion is its failure to take into account normal sadness responses to any losses other than the death of a loved one. It would have been easy to generalize the bereavement exclusion clause (and its accompanying exclusion-to-the-exclusion criteria) to cover all severe losses, but this opportunity was foregone, for reasons explored earlier. Consequently, this constructive attempt to validly delineate the normally sad from the disordered is too limited to adequately address the glaring weaknesses in the symptomatic criteria.

Exclusion for General Medical and Substance-Use-Induced Depressions

The exclusion from MDD diagnosis of depressive conditions that directly result from the physiological effects of medical conditions or substance use simply shifts such cases into alternative disorder categories of Mood Disorder Due to General Medical Condition or Substance-Induced Mood Disorder. These categories, although not our focus here, are subject to their own potential confusions. For example, such disorders are sometimes confused with normal sadness responses to having a medical condition or with sadness in response to the problems that result from using or being addicted to a substance. This is an instance of the complex challenges practitioners face in separating symptoms that indicate depression from similar symptoms that are not disordered or that are the result of different disorders.

The Clinical-Significance Requirement

Perhaps the most important attempt in the *DSM*'s exclusion clauses to distinguish disordered from normal sadness responses is the "clinical significance" criterion, which requires that "the symptoms cause clinically significant distress or impairment in social, occupational, or other important areas of functioning." This clause implicitly acknowledges that even nonbereaved cases that satisfy the duration and symptom criteria might still not involve disorder. However, the clause does not address the basic validity problems of the MDD criteria. It was meant to ensure that the negative consequences of a condition exceed a threshold of significance if the condition is to be clinically relevant and thus potentially classifiable as a disorder,

and it does this successfully. But it does not recognize some crucial distinctions. First, periods of sadness in general, whether normal or disordered, inherently entail negative emotions that involve distress. Indeed, it is hard to imagine having five of the specified symptoms without experiencing distress.

Second, intense normal loss responses almost always involve impairment and diminished interest and ability in various areas of functioning; the very prototype of these responses involves social withdrawal and wanting to be left alone (e.g., one does not feel like seeing friends or going to work). Indeed, intense normal loss responses may be designed to cause distress and social withdrawal to enable one to avoid threats and reconsider one's life and goal structure (see chapter 2).[4] Thus the clinical-significance exclusion might eliminate from the disorder category a few conditions whose feeble symptoms occasion no harm. But it is likely to be used quite rarely because the listed symptoms themselves already involve obvious forms of distress and impairment, rendering the requirement of distress or impairment virtually redundant.[5]

The clinical-significance criterion fails to resolve the problem of distinguishing normal from disordered conditions that satisfy *DSM* criteria because, like the symptom and duration criteria, it potentially applies to both kinds of conditions and fails to address the question of dysfunction. Nor is the addition of the qualifier "clinically significant" helpful in making the distinction clearer because the qualifier is left undefined. Thus the phrase can mean only "significant enough to indicate a clinical—that is, disordered—condition," making the criterion circular with respect to distinguishing normal from disordered conditions.

Implications of the DSM's Own Definition of Mental Disorder

Interestingly, our claim that there is a flaw in the *DSM's* definition of Major Depressive Disorder with respect to distinguishing disordered from normal sadness appears to be implicit in the text of the *DSM* itself. The *DSM's* preface contains a brief general definition of mental disorder that is supposed to be used to determine which conditions are allowed into the manual in the first place. The *DSM-IV's* definition of mental disorder reads as follows:

> In *DSM-IV*, each of the mental disorders is conceptualized as a clinically significant behavioral or psychological syndrome or pattern that occurs in an individual and that typically is associated with present distress (e.g., a painful symptom) or disability (i.e., impairment in one or more important areas of functioning) or with a significantly increased risk of suffering death, pain, disability, or an important loss of freedom. In addition, this syndrome or pattern must not be merely an expectable and culturally sanctioned response to a particular event, for example, the death of a loved one. Whatever its original cause, it must currently be considered a

manifestation of *a behavioral*
individual. Neither deviant be
conflicts that are primarily b
disorders unless the deviance
the individual, as described a

This definition commendably
conditions in terms of the prese
way, without attempting to expl
is a general definition of the con
manual's categories, it follows t
lar disorders presumably shoulc
caused symptoms should count
later editions of the manual ever
nostic criteria with the general
nate because it appears that in n
the specific diagnostic criteria se

The definition of mental disc
symptoms that emerge because
because of socially expectable or
important respects to the "harmf
the background for our discussic
indicate that even conditions tha
orders, because the presence of
result from a dysfunction. The d
cannot be considered a disorder s
desirability, even if there is distre
Rather, the condition is a disorde
symptoms. But, according to this
to external stressful events in the
emotional and other reactions of
disorder partly describes, does no
disorder. Consequently, the *DSM's*
most plausible account of "dysfu
that the criteria for MDD are inva
rally selected loss responses as dis

The Precedent of Co

It may seem impossible that the
agnostic criteria in the *DSM* coul
but also inconsistent with *DSM's*

ecognize some crucial distinctions.
ormal or disordered, inherently en-
ıdeed, it is hard to imagine having
ncing distress.
lmost always involve impairment
ıus areas of functioning; the very
withdrawal and wanting to be left
s or going to work). Indeed, intense
use distress and social withdrawal
r one's life and goal structure (see
ɛclusion might eliminate from the
ɛble symptoms occasion no harm.
the listed symptoms themselves al-
npairment, rendering the require-
ındant.[5]
'esolve the problem of distinguish-
ıtisfy *DSM* criteria because, like the
applies to both kinds of conditions
ɔn. Nor is the addition of the quali-
the distinction clearer because the
ı mean only "significant enough to
lition," making the criterion circu-
n disordered conditions.

the *DSM's* definition of Major De-
ıg disordered from normal sadness
itself. The *DSM's* preface contains
that is supposed to be used to de-
.he manual in the first place. The
as follows:

.s conceptualized as a clinically
ndrome or pattern that occurs
iated with present distress (e.g.,
airment in one or more impor-
ɪcantly increased risk of suffer-
ıt loss of freedom. In addition,
ɔrely an expectable and culturr-
ent, for example, the death of a
must currently be considered a

manifestation of *a behavioral, psychological, or biological dysfunction in the individual.* Neither deviant behavior (e.g., political, religious, or sexual) nor conflicts that are primarily between the individual and society are mental disorders unless the deviance or conflict is a symptom of a dysfunction in the individual, as described above.[6]

This definition commendably distinguishes disordered from nondisordered conditions in terms of the presence of internal dysfunction, albeit in a cursory way, without attempting to explain the concept of dysfunction. Given that this is a general definition of the concept of disorder that should apply to each of the manual's categories, it follows that the sets of diagnostic criteria for particular disorders presumably should meet the general rule that only dysfunction-caused symptoms should count as disorders. However, neither the *DSM-III* nor later editions of the manual ever made a systematic attempt to rectify the diagnostic criteria with the general definition of mental disorders. This is unfortunate because it appears that in many instances the definition is more valid than the specific diagnostic criteria sets are.

The definition of mental disorder, which relies on the distinction between symptoms that emerge because of a dysfunction in the individual rather than because of socially expectable or undesirable conditions, is quite similar in some important respects to the "harmful dysfunction" account of disorder that forms the background for our discussion.[7] In particular, the *DSM* definition seems to indicate that even conditions that manifest certain symptoms may not be disorders, because the presence of a disorder depends on whether the symptoms result from a dysfunction. The definition also usefully asserts that a condition cannot be considered a disorder sheerly on the basis of its personal or social undesirability, even if there is distress or impairment or other harmful symptoms. Rather, the condition is a disorder only if a *dysfunction in the person* causes the symptoms. But, according to this definition, it would seem that a person reacting to external stressful events in the way we naturally react. namely, with certain emotional and other reactions of the kind the *DSM*'s symptom list for depressive disorder partly describes, does not have a dysfunction and thus does not have a disorder. Consequently, the *DSM*'s own definition of disorder, combined with the most plausible account of "dysfunction" as failure of natural function, implies that the criteria for MDD are invalid because they misclassify intense but naturally selected loss responses as disorders.

The Precedent of Conduct Disorder

It may seem impossible that the expert diagnosticians who formulated the diagnostic criteria in the *DSM* could arrive at criteria that are not only invalid but also inconsistent with *DSM*'s own stated definition of disorder. However,

clinical diagnosis is a quite different task from conceptual analysis of the defining criteria that separate disorder from normality. The two require different skills (just as, for example, recognizing chairs when you see them is very different from formulating a principled definition of the concept "chair" that picks out all and only chairs), and it is thus possible for such errors to enter into the manual. Consider an acknowledged precedent: the *DSM-IV* text itself states that the criteria for an important disorder of childhood and adolescence, Conduct Disorder (i.e., a disorder of antisocial behavior, diagnosed by three or more out of a list of behaviors such as theft, running away, etc.), are invalid and encompass some conditions that should not be diagnosed as disorders despite their satisfying the diagnostic criteria. The problem, the *DSM-IV* informs us, is that the symptomatic antisocial behaviors used to diagnose Conduct Disorder may occur in some conditions that are not due to a psychological dysfunction but only to a normal reaction to difficult environmental circumstances.

Here is what the *DSM-IV* has to say about its own Conduct Disorder criteria:

> Concerns have been raised that the Conduct Disorder diagnosis may at times be misapplied to individuals in settings where patterns of undesirable behavior are sometimes viewed as protective (e.g., threatening, impoverished, high-crime). Consistent with the *DSM-IV* definition of mental disorder, the Conduct Disorder diagnosis should be applied only when the behavior in question is symptomatic of an underlying dysfunction within the individual and not simply a reaction to the immediate social context. Moreover, immigrant youth from war-ravaged countries who have a history of aggressive behaviors that may have been necessary for their survival in that context would not necessarily warrant a diagnosis of Conduct Disorder. It may be helpful for the clinician to consider the social and economic context in which the undesirable behaviors have occurred.[8]

This passage says that the *DSM* criteria for Conduct Disorder are not valid when applied to symptoms that could occur as a normal response to circumstances, as, for example, when psychiatrically normal youths join gangs for self-protection in a threatening neighborhood and engage in antisocial behavior as part of required gang activities. Thus the Conduct Disorder criteria do not always pick out dysfunctions. We are making exactly the same point about the criteria for MDD. The symptomatic criteria do sometimes pick out dysfunctions and thus disorders, but they also pick out a potentially large range of normal responses to problematic environments. As in Conduct Disorder, the problem is not particularly hard to see once one considers obvious examples. Yet it is a profound problem that throws into doubt the meaning of much recent research on depression, as we show in later chapters.

In addition, the criteria for Conduct Disorder contain the same kind of "clinical significance" requirement that appears in the MDD criteria. But the textual

comment just quoted implies tha
of adequately distinguishing no
the addition of the clinical-signifi
whether there is sufficient harm f
whether a dysfunction causes the
though the clinical-significance c
too mild to constitute a disorder, t
remains about whether or not the
thus about whether the symptom
circumstances. Precisely the same
inclusion of the clinical-significan

How the *DSM* Attemp Contextual Triggers o

Even if the MDD criteria taken in i
some would answer our criticisms
as a whole. They could say that o
other complementary categories o
address the issue of normal loss re
other categories and features that
We argue that, far from compensa
these complementary categories a
at all or in some cases actually mak
ening the scope of normal sadness

Textual Mention of Norm

Textual commentary that accompa
Disorders does indeed mention the
from depressive disorder. However
forces the problems noted earlier.
after a lengthy discussion that sugg
nated from various other mental di
simply repeats the requirements sta
sion clause), the *DSM-IV-TR* says t

> Finally, periods of sadness are ir
> These periods should not be dia
> less criteria are met for severity
> (i.e., most of the day, nearly eve
> significant distress or impairmer

comment just quoted implies that the Conduct Disorder criteria are incapable of adequately distinguishing normal from disordered conditions even with the addition of the clinical-significance clause. This clause certainly addresses whether there is sufficient harm for a disorder diagnosis, but it does not address whether a dysfunction causes the harm. In the case of Conduct Disorder, even though the clinical-significance criterion eliminates conditions with symptoms too mild to constitute a disorder, the *DSM-IV* recognizes that a separate question remains about whether or not there is a dysfunction causing the symptoms and thus about whether the symptoms represent a disorder or a normal reaction to circumstances. Precisely the same issue remains in the case of MDD despite the inclusion of the clinical-significance criterion.

How the *DSM* Attempts to Address Contextual Triggers of Sadness

Even if the MDD criteria taken in isolation have the problems we have identified, some would answer our criticisms by suggesting that the *DSM* must be looked at as a whole. They could say that our objections to the criteria are dealt with via other complementary categories or other features of the manual that somehow address the issue of normal loss responses. So, in this part, we consider various other categories and features that the *DSM* uses to handle depressive symptoms. We argue that, far from compensating for the weaknesses in the MDD criteria, these complementary categories and features either do not address the problem at all or in some cases actually make things considerably worse by further broadening the scope of normal sadness responses that can be labeled as pathological.

Textual Mention of Normal Sadness

Textual commentary that accompanies the criteria for MDD and the other Mood Disorders does indeed mention the challenge of distinguishing normal sadness from depressive disorder. However, the way it addresses the issue simply reinforces the problems noted earlier. Under a section on "differential diagnosis," after a lengthy discussion that suggests how depressive disorder can be discriminated from various other mental disorders and from bereavement (here, the text simply repeats the requirements stated in the MDD criteria's bereavement exclusion clause), the *DSM-IV-TR* says the following:

> Finally, periods of sadness are inherent aspects of the human experience. These periods should not be diagnosed as a Major Depressive Episode unless criteria are met for severity (i.e., five out of nine symptoms), duration (i.e., most of the day, nearly every day for at least 2 weeks), and clinically significant distress or impairment.[9]

This passage just reiterates the diagnostic criteria for MDD and reasserts that they are sufficient for disorder. The clear implication is that normal periods of sadness never satisfy the criteria. But, as demonstrated earlier in this book, this is not so. Part of the range of normal variation in sadness, especially in response to severe losses and threats, can easily meet *DSM* criteria. Thus, in stark contrast to the textual comment that accompanies the criteria for Conduct Disorder, the MDD comment seems a half-hearted gesture toward acknowledging the problem of distinguishing depressive disorder from normal sadness; however, it only repeats the original error in the criteria.

Multiaxial System

A second way in which the *DSM* tries to address the issue of development of symptoms in response to stressors is via its multiaxial system of diagnosis. This system rates patients on five distinct dimensions that go beyond the diagnostic criteria. Diagnoses of MDD (and all other mental disorders) are recorded on Axis I, personality disorders on Axis II, general medical conditions on Axis III, psychosocial and environmental problems on Axis IV, and global assessment of functioning on Axis V. The various axes are intended to give the clinician a more comprehensive picture of the context of the patient's problem than the diagnostic criteria alone provide. In particular, Axis IV involves reporting psychosocial and environmental problems that affect the diagnosis, treatment, and prognosis of mental disorders and would include stressors that trigger a loss response.

The problem is that the Axis IV grouping of psychosocial stressors simply places them on a completely separate dimension from the diagnoses of disorders. Symptoms that meet criteria for MDD would have *already* been defined as disordered before Axis IV would come into play. This added information, valuable as it may be, does not in any way address the normal-versus-disordered relation between existing stressors and symptomatic responses, and so it fails to address the problem of whether the condition is a psychological dysfunction or a nondisordered response to a stressor. This axis provides a way for clinicians to take stressors into account in case descriptions, not a means of separating disordered from nondisordered conditions that meet symptomatic criteria.

V Codes for Nondisordered Conditions

Third, the *DSM* contains a short section called "Additional Conditions That May Be a Focus of Clinical Attention," which includes nondisordered conditions for which patients often consult professionals. These categories are often called "V codes" after the letter that precedes their numerical diagnostic codes in the *DSM-III*. Among the V codes is Bereavement, under which it is noted that "As part of their reaction to the loss, some grieving individuals present with symptoms characteristic of a Major Depressive Episode. . . . The diagnosis of Major

Depressive Episode is generally r
2 months after the loss."[10] The c
nizes that a condition can satisfy
Depressive Episode and yet not b
ited to grief after loss of a loved o
nondisordered conditions goes, it
clause in the MDD criteria alread

Among the other V codes are s
identity, spiritual, acculturation, a
provide any symptom criteria for
they state only that a condition ca
due to a mental disorder." Thus th
in living are not mental disorders.
symptoms of mental disorders fro
living. In particular, it makes no pr
classify a condition that satisfies th
DSM states that to qualify as a V co
disorder," it in effect means that th
mental disorder, including MDD. C
atic criteria for MDD must be giver
V code. Only residual conditions t
placed under a V code. Therefore,
normal loss responses that satisfy th
section is exactly where many poter

Adjustment Disorder

The main way the *DSM-IV* address
ors is via the diagnostic category of
This category in effect attempts to de
"reactive" depressions that occur in
in formulating such a definition is t
to circumstances, so the criteria mus
mal reactions. The criteria for Adjust
ever, fail to surmount this challenge
gize (i.e., incorrectly treat as disorde
responses beyond those that would fa

Intended to distinguish pathologic
actions, the overall category of Adju
categories, each of which involves a
stressor, including depressed mood,
toms, and a catchall "unspecified" cat
drawal, work inhibitions, and other pr

Depressive Episode is generally not given unless the symptoms are still present 2 months after the loss."[10] The category of Bereavement thus explicitly recognizes that a condition can satisfy the full set of symptomatic criteria for a Major Depressive Episode and yet not be a mental disorder. But this category is limited to grief after loss of a loved one, and consequently, insofar as recognition of nondisordered conditions goes, it just repeats what the bereavement exclusion clause in the MDD criteria already contains.

Among the other V codes are separate categories for academic, occupational, identity, spiritual, acculturation, and phase-of-life problems. The V codes do not provide any symptom criteria for such nondisordered problems; in each case, they state only that a condition can be classified under the category if it is "not due to a mental disorder." Thus the *DSM-IV* does recognize that many problems in living are not mental disorders. However, it gives no criteria for distinguishing symptoms of mental disorders from those that are nondisordered problems in living. In particular, it makes no provision for overriding the criteria for MDD to classify a condition that satisfies those criteria as a normal response. When the *DSM* states that to qualify as a V code the condition must be "not due to a mental disorder," it in effect means that the condition cannot satisfy *DSM* criteria for a mental disorder, including MDD. Consequently, conditions that meet symptomatic criteria for MDD must be given a specific diagnosis as a disorder and not a V code. Only residual conditions that do not satisfy disorder criteria may be placed under a V code. Therefore, the V codes do not address the problem of normal loss responses that satisfy the *DSM* criteria for MDD. In fact, the V-codes section is exactly where many potential diagnoses of MDD likely belong.

Adjustment Disorder

The main way the *DSM-IV* addresses the issue of sadness responses to stressors is via the diagnostic category of Adjustment Disorder With Depressed Mood. This category in effect attempts to define what the *DSM-I* and *DSM-II* used to call "reactive" depressions that occur in response to circumstances. The challenge in formulating such a definition is that most normal sadness is also "reactive" to circumstances, so the criteria must somehow distinguish disordered from normal reactions. The criteria for Adjustment Disorder With Depressed Mood, however, fail to surmount this challenge and thus inadvertently manage to pathologize (i.e., incorrectly treat as disorder) a vast range of additional normal loss responses beyond those that would fall under the criteria for MDD.

Intended to distinguish pathological overreactions to stress from normal reactions, the overall category of Adjustment Disorder encompasses a set of subcategories, each of which involves a specific kind of symptomatic reaction to a stressor, including depressed mood, anxiety, antisocial conduct, mixed symptoms, and a catchall "unspecified" category for physical complaints, social withdrawal, work inhibitions, and other problematic reactions to stress. Adjustment

disorder is a residual "category that should not be used if the disturbance meets the criteria for another specific Axis I disorder," such as MDD.[11]

To qualify specifically as Adjustment Disorder "With Depressed Mood," the condition must meet the general criteria for Adjustment Disorder (discussed next) and, in addition, fulfill the following symptomatic criterion: "This subtype should be used when the predominant manifestations are symptoms such as depressed mood, tearfulness, or feelings of hopelessness."[12] The requirement that any of these symptoms be present is so weak that virtually any normal sadness response would satisfy it. Indeed, in principle, the vague depressive symptom criterion allows diagnosis with just one common sadness-response symptom, such as depressed mood or crying.

However, diagnosis also requires satisfying the general Adjustment Disorder criteria, and the validity of Adjustment Disorder With Depressed Mood thus hangs on these general criteria, which are as follows:

A. The development of emotional or behavioral symptoms in response to an identifiable stressor(s) occurring within 3 months of the onset of the stressor(s).
B. These symptoms or behaviors are clinically significant as evidenced by either of the following:
   1. Marked distress that is in excess of what would be expected from exposure to the stressor
   2. Significant impairment in social or occupational (academic) functioning
C. The stress-related disturbance does not meet the criteria for another specific Axis I or II disorder.
D. The symptoms do not represent Bereavement.
E. Once the stressor (or its consequences) has terminated, the symptoms do not persist for more than an additional 6 months.[13]

Adjustment Disorder, unlike MDD, is specifically limited to conditions that are reactions to triggering events. Clause C formalizes the "residual" character of the diagnosis and implies that Adjustment Disorder With Depressed Mood can be diagnosed only if the individual does not satisfy criteria for MDD. As in the criteria for MDD, the only exemption from disorder status is bereavement; reactions to any other losses that satisfy the criteria are considered disordered.

It is certainly true that the process of adjusting to stressors, or "coping," can go awry and become pathological. The critical issue is whether the criteria for Adjustment Disorder succeed in their intended purpose of distinguishing such disordered reactions from normal-range but intense coping responses that can accompany stressful events.

The criteria require that the symptom(s) must occur within 3 months of the stressor and must end within 6 months of the termination of the stressor. These

timing criteria are designed to en
to a stressor and not independent
normal loss responses are charac
stressor that triggers them; they
stressor and to subside soon after
ments potentially encompass the v
do not distinguish disordered from
ment that the reaction cease with
its consequences) is of particular c
a reaction might be considered pa
after the stressor ceases but takes c

The temporal requirements asi
Disorder criteria make between nc
tirely to whether the condition sa
cal significance" criteria under cri
action must include either "marke
expected from exposure to the stre
occupational (academic) functioni

Regarding the "excess distress"
are inherently prone to be distress
normal responses are (by the princi
capable of being, marked. So this c
disordered reactions comes down to
conditions is "in excess of what wou
is how to construe this criterion. It
symptoms are "in excess of what is e
raises the question of how these cr
from disorder. One obvious alternati
able" as a statistical requirement. F
allow the top half or third (say) of
classified as disordered. But having
reaction does not necessarily imply tl
ample: (1) the individual's meaning s
more problematic or threatening tha
exist within a problematic environm
more enduring than usual; (3) the ir
cultural background or family than o
temperamentally respond more inten

A more charitable interpretation
DSM means whatever is a "proportio
ing the nature and context of the stre
tural meanings of the stressor, are t
rough proportionality is one of the e

ot be used if the disturbance meets
r," such as MDD.[11]
)rder "With Depressed Mood," the
r Adjustment Disorder (discussed
nptomatic criterion: "This subtype
:stations are symptoms such as de-
elessness."[12] The requirement that
that virtually any normal sadness
the vague depressive symptom cri-
ı sadness-response symptom, such

g the general Adjustment Disorder
order With Depressed Mood thus
follows:

.vioral symptoms in response to
ithin 3 months of the onset of

:ally significant as evidenced by

' what would be expected from

or occupational (academic)

)t meet the criteria for another

/ement.
has terminated, the symptoms
nal 6 months.[13]

:ifically limited to conditions that
)rmalizes the "residual" character
ıt Disorder With Depressed Mood
not satisfy criteria for MDD. As in
m disorder status is bereavement;
riteria are considered disordered.
sting to stressors, or "coping," can
al issue is whether the criteria for
ed purpose of distinguishing such
intense coping responses that can

ıust occur within 3 months of the
termination of the stressor. These

timing criteria are designed to ensure that the symptoms are indeed a reaction to a stressor and not independent of it. The problem is that the vast majority of normal loss responses are characterized by a close temporal relationship to the stressor that triggers them; they tend to start soon after the occurrence of the stressor and to subside soon after the stressor abates. Thus the timing requirements potentially encompass the vast majority of episodes of normal sadness and do not distinguish disordered from nondisordered reactions to loss. The requirement that the reaction cease within 6 months of termination of the stressor (or its consequences) is of particular concern because one of the best indicators that a reaction might be considered pathological is that it does *not* gradually subside after the stressor ceases but takes on a life of its own independent of events.

The temporal requirements aside, the distinction that the *DSM* Adjustment Disorder criteria make between normal and disordered coping comes down entirely to whether the condition satisfies at least one of the two specified "clinical significance" criteria under criterion B. To be classified as a disorder, the reaction must include either "marked distress that is in excess of what would be expected from exposure to the stressor" or "significant impairment in social or occupational (academic) functioning."

Regarding the "excess distress" criterion, even normal reactions to stressors are inherently prone to be distressing, and when the stressor is a marked one, normal responses are (by the principle of proportionality) prone to be, or at least capable of being, marked. So this criterion's ability to distinguish normal from disordered reactions comes down to its requirement that the distress in disordered conditions is "in excess of what would be expected" for that stressor. The problem is how to construe this criterion. It cannot be understood as requiring that the symptoms are "in excess of what is expectable in a normal reaction," because that raises the question of how these criteria are supposed to distinguish normality from disorder. One obvious alternative is to construe "in excess of what is expectable" as a statistical requirement. However, the statistical interpretation would allow the top half or third (say) of the distribution of normal responders to be classified as disordered. But having greater than the typical or expected level of reaction does not necessarily imply that one's reaction is due to a disorder. For example: (1) the individual's meaning system and values may make a stressor much more problematic or threatening than it is for most people; (2) the individual may exist within a problematic environment in which the stressor is more serious or more enduring than usual; (3) the individual may come from a more expressive cultural background or family than other individuals do; or (4) the individual may temperamentally respond more intensely than most people do to life events.

A more charitable interpretation is that, by an "expectable" response, the *DSM* means whatever is a "proportionate" response when all the factors, including the nature and context of the stressor itself, as well as the subjective and cultural meanings of the stressor, are taken into account. We argued earlier that rough proportionality is one of the earmarks of a nondisordered loss response.

If the first component of criterion B is interpreted as specifying that a reaction "in excess" is outside the range of proportional responses, then, taken by itself, it is potentially a valid indicator of dysfunction and does correctly place some disorders into the Adjustment Disorder category and avoid obvious false positives.

But then there is the problem of criterion B's second component, impairment in social or occupational functioning. This, by itself, is offered as a sufficient alternative for classifying a condition as disordered. Unfortunately, it fails to exclude great numbers of normal loss response conditions. Whenever major stressors occur, it is likely that people will suffer impairment in their social, occupational, or academic functioning. Just the time and concentration it takes to deal with the stressor, the emotional feelings that make it difficult to focus on routine tasks, and the real-life changes that people must make can easily lead them to resist usual tasks and roles. Moreover, the issues and challenges that major stressors trigger may make some role functioning seem temporarily insignificant by comparison, causing a loss of motivation and interest. Virtually any low mood might have such consequences. Thus, even if the "marked distress" criterion is charitably interpreted, the flaws in the alternative impairment criterion ensure that a vast number of normal loss responses can be diagnosed as Adjustment Disorders.

We conclude that the criteria for Adjustment Disorder and for its subtype Adjustment Disorder With Depressed Mood potentially classify as disordered an enormous number of normal responses that are triggered by stressors and that subside after the stressor ends, just as such responses are designed to do. And they do so on the basis of as little as one symptom that reduces role functioning. Indeed, any normal loss response of any consequence that does not fall under the *DSM* criteria for MDD is almost sure to fall under the criteria for Adjustment Disorder With Depressed Mood.

The flaws in the Adjustment Disorder category are so apparent that researchers and epidemiologists have largely ignored it. They have clearly "voted with their feet" that Adjustment Disorder is not of interest, judging from the very low numbers of research studies on it and the lack of growth in those numbers, which stand in stark contrast to the growth of research on other *DSM* categories in general and on MDD in particular. In 1980, 80 medical articles contained "adjustment disorder" in their titles, a number that actually declined to 55 articles in 2005.[14] By the latter year, nearly 158 articles appeared with "depression" in their titles for each article about adjustment disorder. In short, MDD, not Adjustment Disorder With Depressed Mood, has become the operative category for the field when it comes to studying depressive states. This neglect of Adjustment Disorder by researchers appears to be justified. The diagnosis suffers from such glaring problems in distinguishing normal from disordered conditions that it has collapsed as a serious target of research under the weight of its own invalidities. However, within the clinical realm, the diagnosis of Adjustment Disorder may nonetheless sometimes still be useful as a way of providing a potentially

reimbursable label for reactions to
be genuine disorders but that often

Other Depression-Rel
and Features of the *L*

Subthreshold Diagnoses I

Conditions that fail to meet the full
that include some symptoms menti
conditions. The *DSM-IV* placed a ne
would subsume such conditions in
vided for Further Study." It would rec
the nine criteria for MDD, as long as
minished interest or pleasure. In oth
and various exclusions, it is essential

As we shall see in the next chapte
propose that subthreshold conditions
of these recommendations seriously a
old conditions opens the floodgates to
sponses that are not even particularly
could encompass virtually all significa
far, however, the *DSM* has not adopted

Subthreshold Diagnoses II: Not Otherwise Specified

Nevertheless, the *DSM* does already
their discretion can classify as depres
do not meet the *DSM* criteria for MI
for many other kinds of categories, tl
basket" category of Mood Disorder N
main purposes of this category is to
that do not meet the criteria for any s

The manual's introduction includ
Specified Categories" that identifies tl
may be appropriate. The first applies t

Enough information available to in
ent, but further specification is no
sufficient information to make a n
clinical features of the disorder do n
cific categories in that class.[17]

reimbursable label for reactions to stressful circumstances that may or may not be genuine disorders but that often deserve and need clinical attention.

Other Depression-Related Categories and Features of the *DSM-IV*

Subthreshold Diagnoses I: Minor Depression

Conditions that fail to meet the full symptomatic or duration criteria for MDD but that include some symptoms mentioned in the criteria are called "subthreshold" conditions. The *DSM-IV* placed a new category, Minor Depressive Disorder, which would subsume such conditions in an appendix on "Criteria Sets and Axes Provided for Further Study." It would require only two, instead of five, symptoms from the nine criteria for MDD, as long as one symptom is either depressed mood or diminished interest or pleasure. In other respects, such as the duration requirement and various exclusions, it is essentially the same as Major Depressive Disorder.[15]

As we shall see in the next chapter, various arguments in the recent literature propose that subthreshold conditions should be defined as genuine disorders. None of these recommendations seriously addresses the problem that allowing subthreshold conditions opens the floodgates to diagnosing as disorders normal sadness responses that are not even particularly intense or enduring. Indeed, such a category could encompass virtually all significant loss responses or periods of sadness. Thus far, however, the *DSM* has not adopted minor depression as an official category.

Subthreshold Diagnoses II: Mood Disorder Not Otherwise Specified

Nevertheless, the *DSM* does already specify that mental health professionals at their discretion can classify as depressive disorders subthreshold conditions that do not meet the *DSM* criteria for MDD. This is due to the fact that, as it does for many other kinds of categories, the manual includes an additional "wastebasket" category of Mood Disorder Not Otherwise Specified (NOS). One of the main purposes of this category is to diagnose "disorders with mood symptoms that do not meet the criteria for any specific mood disorder."[16]

The manual's introduction includes a section titled "Use of Not Otherwise Specified Categories" that identifies the situations in which an NOS diagnosis may be appropriate. The first applies to conditions for which there is

> Enough information available to indicate the class of disorder that is present, but further specification is not possible, either because there is not sufficient information to make a more specific diagnosis or because the clinical features of the disorder do not meet the criteria for any of the specific categories in that class.[17]

The intention here was no doubt the legitimate one of giving clinicians the flexibility to diagnose occasionally clear disorders that do not quite meet the official threshold for a more specifically named condition in a class. But applying the NOS category to depressive disorder, with no precautions about distinguishing it from normal reactions, could allow clinicians to diagnose as disorders many normal responses that are not intense enough to meet the five-symptom, 2-week threshold.

The second situation the manual specifies as one in which the NOS category can be used is when "the presentation conforms to a symptom pattern that has not been included in the *DSM* classification but that causes clinically significant distress or impairment."[18] This is equally problematic because both normal and disordered sadness can easily possess significant distress and role impairment. So, when it comes to loss responses, the Mood Disorder NOS category in effect gives clinicians carte blanche to classify normal reactions as disorders.

Dysthymic Disorder

A second category of depressive disorders in the *DSM-IV* is Dysthymic Disorder. Conceived in part as a concession to psychodynamic clinicians, this disorder was substituted for the traditional category of neurotic depression (and actually appeared under the title "Dysthymic Disorder (or Depressive Neurosis)" in the *DSM-III*).[19] Its criteria are quite different, however, from those for traditional neurotic depressions, which included excessive but often time-limited reactions to specific stressors. Diagnosis of Dysthymic Disorder requires a disturbance of mood and only two additional symptoms, but it also requires that the symptoms must have lasted for at least 2 years (1 year for children and adolescents) and during that time must have been present for most of the day on most days. Like MDD, Dysthymic Disorder is diagnosed solely on the basis of symptoms, without reference to such factors as chronic stressors (e.g., the gradual decline and death of an ill child) that might distinguish normal from disordered states of chronic depressive symptoms. Nor do the symptomatic criteria allow a distinction between depressive disorder and normal-range melancholic personality or temperament, the latter identified since the time of Aristotle. These problems present major challenges for the validity of the Dysthymic Disorder category itself, and certainly its inclusion as a category of milder but chronic depressive conditions does nothing to fix the problems that stem from the lack of adequate distinction between normality and disorder in the MDD criteria.

Melancholic Major Depressive Disorder

For some persons who meet the MDD criteria, the *DSM* specifies a "With Melancholic Features" subcategory. This classifies an individual who either has lost pleasure in all or almost all activities or who does not react to usually pleasurable stimuli and who displays three additional symptoms from a list that includes a distinct quality of the depressed mood in contrast to usual sadness,

greater severity in the morning, ea
tor retardation, weight loss, and ex

The subcategory of melancholi
cases of *endogenous depression*, wh
instances of depressive disorder.[20]
term *endogenous* because, by traditi
etative" or seemingly physiologicall
triggering circumstances, all of wh
etiological assumption. Instead, th
melancholic subcategory. Conseque
tomatic criteria classify as melanch
and would not traditionally be cons

It is possible that, due to their sp
ancholic depressions may, on averag
types of depression. But melancholi
of *DSM* Major Depressive Disorders
and other depressive conditions car
lems of the overall MDD criteria.[21]

Would it have helped to resolve t
disordered sadness if the *DSM* had
pression to reflect the traditional no
sive" depressions? As we argued in ch
does not adequately distinguish disor
dogenous depressions are generally
by virtue of a disproportionate symp
triggering loss. The *DSM* justly aband
not find an adequate replacement for

Conclusion

The symptom-based diagnoses in th
proved previous efforts to classify de
ambiguous definitions of depression
them sets enhanced communication a
diagnosing of depression. Researchers
kinds of participants, and clinical diag
the same types of conditions.

These undoubted advances, howev
the symptom-based diagnoses did not
with the presence of disorder from exp
explanatory features of the *DSM* that dea
the problem. The manual's own defin

greater severity in the morning, early-morning awakening, marked psychomotor retardation, weight loss, and excessive guilt.

The subcategory of melancholia was intended to correspond to traditional cases of *endogenous depression*, which were considered to be particularly clear instances of depressive disorder.[20] However, the *DSM* does not actually use the term *endogenous* because, by tradition, that term connotes certain types of "vegetative" or seemingly physiologically based symptoms and a lack of any external triggering circumstances, all of which do indeed suggest disorder but involve an etiological assumption. Instead, the *DSM* uses symptoms alone to diagnose the melancholic subcategory. Consequently, many conditions that the *DSM*'s symptomatic criteria classify as melancholic do have associated precipitating stresses and would not traditionally be considered "endogenous."

It is possible that, due to their special symptomatic requirements, *DSM* melancholic depressions may, on average, be actual disorders more often than other types of depression. But melancholic depressions make up only a small fraction of *DSM* Major Depressive Disorders. Thus the distinction between melancholic and other depressive conditions cannot yield any solution to the validity problems of the overall MDD criteria.[21]

Would it have helped to resolve the problem of distinguishing normal from disordered sadness if the *DSM* had formulated the criteria for melancholic depression to reflect the traditional notion of "endogenous" in contrast to "reactive" depressions? As we argued in chapter 1, the endogenous-reactive distinction does not adequately distinguish disorder from nondisorder because, although endogenous depressions are generally disorders, so are many reactive depressions by virtue of a disproportionate symptomatic response to the magnitude of the triggering loss. The *DSM* justly abandoned this distinction but, unfortunately, did not find an adequate replacement for it.

Conclusion

The symptom-based diagnoses in the *DSM-III* and *DSM-IV* in many ways improved previous efforts to classify depression. They overcame the cursory and ambiguous definitions of depression found in previous manuals. Explicit criterion sets enhanced communication among researchers and clinicians about the meaning of depression. Researchers could create more homogeneous populations of participants, and clinical diagnoses had greater chances of referring to the same types of conditions.

These undoubted advances, however, also had costs. The main cost was that the symptom-based diagnoses did not validly distinguish depressions that indicate the presence of disorder from expectable reactions to situational contexts. The many features of the *DSM* that deal with responses to stressors fail to resolve this problem. The manual's own definition of mental disorders, combined with

the empirical data cited in chapter 2, suggests that its criteria for depressive disorder are not valid. The multiaxial system does not help because it uses the relevant axis of psychosocial stressors only to supplement, not to modify, a diagnosis of disorder. The category of Adjustment Disorder merely compounds the problem because it pathologizes even those normal reactions that display fewer than the usual symptoms and that go away when the stressor ceases. Nor does the inclusion of V codes overcome the fundamental problem that all conditions that meet diagnostic criteria must be diagnosed as disorders. It would have been easy enough for the definition of MDD to have included a more extensive set of exclusion criteria comparable to the exclusion for bereavement, but this was not attempted. The result is a major invalidity that leads to the pathologization of intense normal sadness.

Kraepelin conceptually embraced the "with cause" versus "without cause" distinction, although it was not an important practical consideration in classifying his inpatient populations. By the time the *DSM-III* was published in 1980, outpatient therapy was much more common, and, consequently, the range of the problems people brought to psychiatrists had enormously expanded. Just when it would have been most useful to further develop the "with cause" versus "without cause" distinction so as to avoid false positive diagnoses, the *DSM-III* abandoned the distinction and thus inadvertently reclassified as mental disorders many conditions that were problems of living. The resulting problems went unremedied in subsequent editions of the manual. But the problem of pathologizing normal sadness does not end there. The next step in transforming normal unhappiness into mental disorder came when the symptom-based logic behind the *DSM-III* and *DSM-IV* went beyond the clinic and formed the basis for studies of depression among untreated individuals in the community.

6

Importing Pathol

The transformation of i
order occurred in seve
symptom-based research criteria fo
and RDC provided an initial step. Su
that had traditionally protected dia
as disorder and thus created the pc
pression was, however, primarily con
severely afflicted community membe
context, the symptom-based criteria
ders from other serious disorders, a
immediately realized. The second ste
symptom-based logic to clinical prac
patient practices and community clir
of mental distress. Applying deconi
group of outpatients made it more lil
to those suffering from normal sadnes

However, several factors work to n
criteria to normal sadness in outpati
reject, so although many individuals d
they seek treatment only after they a
lems and not to stressful situations.[1] M
entivements by insurance reimburser
possible, can still use their commonse
DSM criteria and to recognize when a
in need of reassurance and support t
transient feelings.

It is not, then, in the clinical cont
normal and abnormal is most in dang
and transformation of ordinary sadne
DSM's criteria, developed primarily fc

1.