Evolutionary foundations for psychiatric diagnosis: making DSM-V valid

Randolph M. Nesse and Eric D. Jackson

The third and fourth editions of the Diagnostic and Statistical Manual of Mental Disorders (DSM) have brought much-needed reliability to psychiatric diagnosis. However, as is often the case, progress comes at a price. In this chapter, we support Wakefield’s argument that DSM-III and DSM-IV typically ignore one of the most fundamental distinctions in medicine— the distinction between symptoms and the situations or diseases that cause them. In the case of emotional disorders, such as mood and anxiety disorders, this mistake is particularly deplorable, because many emotions are responses that evolved because they are protective in untoward circumstances. Here we suggest that an evolutionary perspective can advance the nosology of emotional disorders in several ways. First, this perspective confirms that the normality of an emotion depends necessarily on the context. Furthermore, it notes that variations in brain mechanisms that make a person susceptible to anxiety or depression are only sometimes diseases; more often they may have the same causal significance as variations in brain mechanisms that make a person especially prone to cough or fever during a cold. An evolutionary perspective also indicates that biologically normal responses may be aversive and even harmful to individuals. Finally, it suggests the importance of a detailed and evolutionarily informed analysis of the motivational structure of every patient’s life.

Hundreds of researchers and clinicians have collaborated for the past three decades to revise the diagnostic criteria for mental disorders (Wilson 1993). The products of their labors are the source of widespread dissatisfaction and apparently irresolvable debates (Beutler and Malik 2002; Horwitz 2002). Clinicians often ignore the official diagnostic system. Researchers find themselves constrained by categories with no theoretical foundation and questionable reliability that include heterogeneous patients who show vast comorbidity. Nonprofessionals and experts look at prevalence rates of 50% and ask if there is a scientific justification for defining what is pathological. Even the architects of the system suggest the need for fundamentally new perspectives:

Science strives for simplicity of explanation. Descriptive models tend to be piecemeal and complicated. We are at the epicycle stage of psychiatry where astronomy was before Copernicus and biology before Darwin. Our inelegant and complex current descriptive system will undoubtedly be replaced by explanatory knowledge that ties together the loose ends. Disparate observations will crystallize into simpler, more elegant models that will enable us not only to understand psychiatric illness more fully but also to alleviate the suffering of our patients more effectively.

(Frances and Egger 2003)

Such extensive dissatisfaction after Herculean efforts suggests that persisting in the same path will not solve the problem. This chapter argues that evolutionary behavioral biology is a crucial but neglected scientific foundation for psychiatric nosology. Posing evolutionary questions about why we all are so vulnerable to negative emotions highlights a fundamental misunderstanding at the heart of the Diagnostic and Statistical Manual of Mental Disorders (DSM) that has kept psychiatric diagnoses artificially different from those in the rest of medicine. In the rest of medicine, symptoms such as pain and cough are carefully distinguished from disease such as appendicitis and pneumonia. In psychiatry, we are often trying to craft diagnoses based on symptoms, with predictable frustration. An evolutionary understanding of emotions reveals why the quest for simple criteria for emotional disorders is so frustrating, and where we can look for solutions.

6.1 Diagnosis and its discontents

Our core argument is simple. Negative emotions can be normal and useful in certain situations, so, except in the extreme cases, distinguishing normal and abnormal emotions requires close attention to the situation. The logic is that of the medical model. Consider pain. Pain is normal when its severity matches the amount of tissue damage. Pain is pathological when it is disproportionate to the cause. Decisions about normality and pathology depend on the situation.
The logical response to this argument would be to modify diagnostic criteria to take situations into careful account. As Jerome Wakefield has suggested, for instance, the grief exclusion for depression could be expanded to include other dire circumstances that can cause normal symptoms of depression (Wakefield et al. 2007). Instead, the DSM-V Committee is now apparently considering eliminating the grief exclusion! Instead of simply being aghast at such obliviousness, we should try to understand this response; it can help us understand the problem: the absence of any theoretical foundation for validity, with a resulting huge over-emphasis on reliability. This is understandable on two counts.

First, allowing exclusions for situations such as having a child with cancer, or loss of a marriage or job, would decrease reliability. Who is to say if a particular situation is severe enough to account for the symptoms? Reliability would decrease, and that would be fatal to many studies whose results are already on the border of significance because of the limited reliability of current criteria.

The second issue is more profound. We have no scientific foundation for establishing the validity of criteria for diagnosing emotional disorders. The foundation is being constructed by those working to describe how emotions evolved, how they give advantages, and how selection shaped the mechanisms that regulate them. This work is, however, just getting under way, and it is revealing the inherent difficulties of diagnosing disorders that result from a dysregulation of protective responses. The challenge is hard enough for physical responses such as pain, fever, and fatigue. For emotional responses, the appropriate intensity depends not only on the objective situation, but on how the individual appraises the meaning of the situation for his or her ability to reach personal goals. Emotions arise not from events; they arise from an individual’s motivational structure, that is, from the interaction of an objective external situation with an individual’s goals, strategies, and subjective assessments of ability to reach these goals and strategies.

Such complex causes, different in each case, make it very difficult to formulate reliable diagnostic criteria. If the decision about whether symptoms are normal or abnormal depends on a decision about the severity of the life situation, subjective judgment is unavoidable. The obvious solution is to ignore the situation and focus entirely on the severity and duration of symptoms. If this strategy was used in internal medicine, “cough disorder” would be diagnosed whenever the frequency, duration, and severity of a cough exceeded defined thresholds, irrespective of the cause of the cough. The problem is, of course, that life situations cannot be measured as objectively as a pulmonary infiltrate. Change will eventually come as researchers discover that their findings become stronger when they differentiate subpopulations according to how disproportionate symptoms are to the situation.
This transition will take time. It will be facilitated by creating methods to measure variables that are hard to measure, such as the size of the gap between a person’s resources and aspirations, the extent to which the problem is an objective inability to get crucial resources, the scale of the individual's aspirations, and the extent of distorted negative thinking. But it will also be sped by neuroscientists and other psychiatric researchers who recognize the opportunity to ground their work in behavioral biology. Perhaps this chapter, and others like it, will fire the curiosity of some researchers to explore our growing knowledge about how evolutionary behavioral biology can inform psychiatric research.

Although this chapter emphasizes the utility of evolutionary principles for classifying emotional disorders, the same principles are also useful for classifying other psychiatric disorders. For instance, behavioral disorders such as addiction or eating disorders make much more sense in an evolutionary framework. Personality disorders can be organized based on the strategies people use to influence other people. Even psychoses and neurological conditions are illuminated by evolutionary considerations of the selection forces that maintain the frequency of predisposing genes and how they interact with novel aspects of the modern environment. Here, however, the focus is on the emotions and the categories that describe their disorders.

Much has been written about how an evolutionary approach can help distinguish pathological from non-pathological conditions (Wakefield 1992; McGuire and Troisi 1998; Clark 1999; Troisi and McGuire 2002). Wakefield's concept of “harmful dysfunction” brings a biological foundation to the question of where normal stops and pathology begins (see Chapter 5). Dysfunction grounds diagnosis in the selective advantages of normally operating brain mechanisms (but see Chapters 7 and 8). By also requiring a condition to be “harmful”, Wakefield’s approach acknowledges that what is good for our genes is not necessarily good for our selves, and what is good in one culture may not be good in another. In more recent work, Horwitz and Wakefield make a powerful case for basing psychiatric diagnosis on an evolutionary understanding of emotions. They point out that if depression symptoms can be normal in bereavement, they very likely can also be normal in other situations, therefore careful consideration of the situation is essential to any scientific nosology for depression (Horwitz and Wakefield 2007).

6.2 From clinical diagnosis to the DSM

The history of mental illness taxonomy began with highly speculative informal categories originating thousands of years ago. At turns biological, phenomenological, and moral in orientation, such informal systems prevailed well into
the nineteenth century, when Emil Kraepelin took the first steps toward modern, systematic classification in collaboration with his colleague Allen R. Diefendorf (Kihlstrom 2002). In the USA formal classification systems for mental disorders were first adopted not by clinicians, but by the federal government because of its need to track asylum populations accurately. This encouraged the American Medico-Psychological Organization (AMPA) to publish the first standardized psychiatric nosology, the *Statistical Manual for the Use of Institutions for the Insane (Statistical Manual)* in 1918. The absence of the word “diagnosis” in the title accurately represents the marginal utility of the manual to the era’s mental health practitioners.

The *Statistical Manual* was revised for the last time in 1942, just as the USA entered World War II (Grob 1991; Houts 2002). Military practitioners found the statistical categories woefully inadequate to describe battlefield psychological casualties. Dr. George Raines, then head of the American Psychiatric Association (APA) Committee on Statistics and Nomenclature, noted in the introduction to the first edition of DSM that “only about 10% of the total cases seen [in World War II] fell into any of the categories ordinarily seen in public mental hospitals” (American Psychiatric Association 1952, p. vi). Such dissatisfaction led the APA to replace the *Statistical Manual* with a new standardized nosology in 1952: the first edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-I).

Although more useful to practitioners, DSM-I and its revision, DSM-II (American Psychiatric Association 1968), were unsatisfactory for research. Prior to the formulation of the Research Diagnostic Criteria and the publication of the third edition of the *Diagnostic and Statistical Manual* (DSM-III) (American Psychiatric Association 1980), research reports were hard to compare because the subjects in one study of “depression” might have quite different conditions from those in another. Critics pointed to such inconsistencies to argue that psychiatry was unscientific or even that mental illnesses were not diseases at all (Szasz 1974). Examples of malingering were published in *Science* as evidence for the subjectivity of psychiatric diagnosis (Rosenhan 1973). At about the same time, the utility of psychotropic drugs was being widely recognized and insurance companies began paying only for the treatment of specific medical disorders. These several crises combined to create a consensus that psychiatry should become more like the rest of medicine (Houts 2002; Jackson unpublished). Operationalizing diagnostic criteria was the obvious place to start.

The committee charged with creating the DSM-III quickly found there would be no agreement on a theoretical foundation for psychiatric nomenclature (Wilson 1993). Psychoanalysts remained powerful, and their views of
mental disorders were fundamentally at odds with "biological" psychiatrists, who emphasized the brain origins of mental disorders. To get past this impasse, the DSM created diagnostic categories avowedly without theoretical foundation. The goal was a system derived empirically from clinical observations of observable signs, symptoms, and the disease course. Building on criteria from the International Coding Diagnoses and the Research Diagnostic Criteria group at Washington University (Feighner et al. 1972), the DSM-III and DSM-IV attempted to create categories defined by observable data (American Psychiatric Association 1994).

The inauguration of operationalized diagnoses transformed psychiatry (Guze 1992; Wilson 1993; Jackson unpublished). Indeed, the history of medicine contains few transitions so sudden and complete (Shorter 1997). Prior to the DSM-III, psychiatrists' diagnostic categories were theoretically based and used to complement highly valued narrative explanations for how an individual came to have his or her particular constellation of symptoms. Clinicians crafted idiosyncratic explanations for a particular individual's problems in much the same way that historians explain the origins of a war or economic collapse in a particular country. Nomothetic (universally applicable) principles were incorporated into such explanations, but different clinicians used different principles. For instance, psychoanalysts emphasized the ubiquitous importance of defenses against Oedipal wishes, while behaviorists emphasized the reinforcement history. Arriving at a diagnostic formulation was an occasion for deep thought, sophisticated discussion, theoretical battles, and frequent flights of fancy. Two diagnosticians often arrived at plausible formulations with little in common and no way to decide between them. Reliability was low. Such diagnoses were nearly worthless for research.

Current criteria are nearly the polar opposite of their predecessors. Individualized explanations for symptom constellations have been replaced by categories defined by the presence or absence of specific signs and symptoms. For instance, a diagnosis of major depression applies to anyone who has had at least five of nine symptoms for at least 2 weeks, at least one of which is depressed mood or lack of pleasure (American Psychiatric Association 1994). Precipitating events are not taken into account, with the exception of bereavement in the past 2 months. Whether symptoms arise during a relaxing vacation or a stay in intensive care is irrelevant. Such exclusion of life context is mindless, but it does sidestep the serious problem of how to measure the kind and severity of precipitants. If criteria for depression required assessing the severity of recent life events, complexity would increase and reliability would plummet.

The quest for criteria that yield reliable diagnoses is well justified. If different clinicians examining a patient arrive at different diagnoses, the system is not all
that useful (Goodwin et al. 1979). Explicit criteria made possible standardized interviews that further enhanced reliability (Spitzer et al. 1992). Versions useable by lay interviewers have made extensive epidemiology possible for the first time, not just in the USA, but in over 39 countries where the same questions are administered using the same instrument translated into different languages (Kessler and Ustun 2004). This is real progress, and the data are useful for public health planning as well as research.

In short, the DSM has been essential for most recent progress in psychiatry. Treatment trials now target groups of well-defined patients and the results can be applied to other similarly defined groups. Research studies can measure genes, neurotransmitters, or brain structures in well-characterized groups of patients as compared to controls. Reliable diagnostic criteria have advanced psychiatric research more than any individual research project could.

6.3 The price of progress

Given such dramatic progress, why such dissatisfaction with the DSM approach to diagnosis? Many objections are based on the tangible factors outlined above—high comorbidity, heterogeneity within groups, and questionable reliability (Beutler and Malik 2002; Phillips et al. 2003; Watson 2005). However, larger issues are even more important.

First, the distinction between normal and abnormal remains fundamentally arbitrary. For cancer, pneumonia, rheumatoid arthritis, and pinworms you either have the condition or you don’t. A zone of rarity separates the condition from normal (Kendell 1975). Most emotional disorders offer no such clean demarcation, leading some to suggest that diagnoses should be dimensional instead of categorical. However, to communicate, humans tend to use words that refer to categories or essences. People demand to know the boundary that separates pathology from normality. Dimensions are not diagnoses. Even high blood pressure is defined by a specific cut-off.

Second, the DSM diagnoses are often presented as products of clinical observation unconnected explicitly to any theory of human behavior. However, this presentation is not quite correct. Because explicit theories are excluded, the DSM criteria tacitly foster thinking about mental disorders as if they are diseases. This makes them fit easily into neuroscience models that seek to identify brain abnormalities correlated with each disorder.

Third, is the problem of how to incorporate context (Faust and Miner 1986). The DSM approach relegates much of what we know about the effects of life events to “stress”, as if stress hormones mediated most adverse effects of social experience. Clinicians understand the far more complex relationships between life events and psychological structures (Brown and Harris 1978; Monroe et al. 1992; Turner and Brown 2000).
Many see the need to adjust the diagnostic threshold depending on the situation, lowering it for apparently unprovoked symptoms, increasing it in extreme life situations. However, with the exception of bereavement, the DSM criteria ignore context (Wakefield and First 2003). For instance, the diagnosis of panic disorder is applied whenever someone has symptoms for a month after recurrent unexpected episodes that include four out of ten possible panic symptoms. It makes no difference whether the patient had the onset in a grocery store or in a prison camp. The reason for this rigidity is that attempts to include context would require difficult-to-define objective criteria for levels of provocation.

At the same time, almost everyone recognizes the need to consider the circumstances in order to judge whether an emotion is normal or not. Following the debut of the multi-axial diagnostic system in DSM-III, it appeared that the editors of DSM had at least partially recognized the need to integrate life circumstances and context. Severity of psychosocial stress (Axis IV) and level of adaptive functioning (Axis V) were added to enrich the clinical context of the individual (Klerman 1984). However, the inclusion of environmental circumstances in separate axes excludes important contextual information from their important role in making Axis I diagnoses.

The DSM gives us categories for emotional disorders, but says nothing about what these disorders are. Are they diseases? Disorders? Are some merely responses to life circumstances? Is the cause located mainly in brain differences, in cognitive habits, or in exposure to environmental events? Almost everyone pays lip service to the bio-psycho-social model, but few are willing to get into the complexity of how individual differences interact with situations, events, and cognitions to give rise to symptoms that have evolutionary significance (Gilbert 1989).

Thinking about patients as DSM diagnoses instead of people impoverishes clinical understanding (Faust and Miner 1986). For instance, a resident recently concluded a case presentation by saying, “The diagnosis is major depression so I prescribed an SSRI.” When asked why this person was depressed now, the resident replied “Well, we think depression is caused mostly by genetic factors, but also by stress,” omitting any mention of why this particular patient was depressed now. When pressed to do so, he explained that there was a family history of depression and the patient had been abused in childhood, was in a bad marriage, and had recently lost his job after a drink-driving conviction. The resident clearly imagined that his job was to place his patient in the category “major depression” and to prescribe a treatment that was usually effective for someone in that category. He had not even tried to figure out whether the person had had a previous satisfactory and stable life.
adaptation, whether the alcohol use initiated the marital problems or came later, and whether or not the person was capable of maintaining good relationships. In short, like many young clinicians, the resident viewed DSM criteria as if they described specific diseases with specific consistent causes. He assumed that the diagnosis contained all that he needed to know to arrive at a treatment plan.

The same physician would never undertake such a crude approach to diagnosing and treating cough or pain. If a patient presented with a severe cough and fever, he would not be satisfied with a diagnosis of “cough disorder”, he would instead consider all the possible causes of cough, and would not prescribe treatment until arriving at the best possible understanding of why this person had this cough now. Is it chronic obstructive pulmonary disease (COPD), or pneumonia, or congestive heart failure, or COPD and congestive heart failure complicated by pneumonia? The physician would find out whether the individual was especially vulnerable because of immunosuppression or steroid use, if there were exposures to infectious agents, and if the person had allergies. General physicians recognize that cough is not a disease, it is a response to a disease. Likewise, while pain can be abnormal, physicians recognize that pain is usually a response to pathology, not a disorder in itself. Psychiatrists sometimes think of anxiety as a potentially useful response to a danger, but other emotions such as depression and jealousy are usually thought of as abnormalities instead of being recognized as potentially useful responses to untoward situations.

### 6.4 The basic fault

The flaw in the DSM approach to emotional disorders is fundamental: the DSM fails to distinguish protective responses from diseases. This flaw is by no means new; the DSM merely extends the Kraepelinian tradition. Kraepelin excluded etiology and anatomic considerations from mental disorder classification because reliable information was not accessible except in the case of obvious injuries and post-mortem assessment of neural lesions (Kihlstrom 2002). In his 1904 textbook, Kraepelin recognized the limits of a nosology based on symptoms, but he also noted that diagnostic systems based on a comprehensive knowledge of symptoms or pathological anatomy or etiology should provide “uniform and standard classifications” that mapped well onto one another, no matter what the starting point was (Kraepelin and Dierdorf 1907).

It is a short leap from this to equating the outcome of exhaustive identification of symptomatology with the exhaustive identification of etiology; if all nosologies carve up the pie identically, then any one system should work as well as any other. This explanation is especially appealing if some systems are inaccessible,
as neural systems were in the late 1800s. However, assuming that symptomatic categories will match etiological categories comes at the high price of blurring the directional relationship between cause and effect, leading to two kinds of errors. First, categories based on symptom constellations may contain subgroups that arise from fundamentally different causes. Second, such categories fail to distinguish symptoms that arise from pathological causes from those that are aroused by normally functioning systems. The former is an error of failing to distinguish distinct disorders (e.g., yellow fever vs spotted fever), while the latter fails to distinguish disorders from the symptoms of disorders (e.g., mistaking fever or cough for disorders, when they are actually protective responses to the disorder of pneumonia).

The rest of medicine long ago replaced symptomatic diagnoses such as “cough disorder” with etiologically based diagnoses such as pneumonia or lung cancer (Kihlstrom 2002). The rest of medicine recognizes cough, fever, pain, nausea, fatigue, diarrhea, vomiting, and inflammation as responses to diseases, not diseases themselves. These responses are aversive, and they can be dangerous, disabling, and even fatal. High fever can cause convulsions and diarrhea causes thousands of deaths each year. Nonetheless, fever, diarrhea, and other defenses are the body’s adaptive responses to problems, not usually diseases themselves. They give important clues to the diagnosis, but they are themselves diagnoses only in special circumstances.

One circumstance is when the cause cannot be found. For instance, “fever of unknown origin” is a stand-in for a diagnosis when no reason for a fever can be identified. The other circumstance is when the system that regulates the response is presumed to be abnormal, as is the case in chronic pain syndromes. Chronic fatigue is likewise usually thought to arise from an abnormal regulation system. When every other possible cause has been eliminated, even fever or pain may be attributed to an abnormal regulation mechanism.

The error of failing to distinguish defenses from diseases needs a name. Most simply it can be called “the fallacy of mistaking defenses for diseases”. It could be called “the DSM fallacy” because the DSM so resolutely ignores this basic medical distinction. The DSM takes great pains to define when symptoms are severe enough to justify a diagnosis, but it mostly ignores the more fundamental distinction between symptoms and the problems that arouse them.

This argument is based on the supposition that negative emotions are protective reactions akin to pain and fever. The next section reviews reasons to think this is correct. However, major differences between physical protective responses and emotional responses make the correspondence hard to see at first. The situations that arouse fever and cough are observable changes in specific tissues. Most arise from diseases or injuries. The situations that arouse negative
6.5 EVOLUTION AND EMOTIONS

Emotions are also adverse, but few are specific diseases with identifiable tissue pathology. Many are injuries to social resources such as relationships or social status, which are less tangible despite their importance to function and Darwinian fitness. Some situations, such as exclusion from a group, directly arouse negative emotion. Other connections between situations and emotions are far less direct, such as the anxiety that follows a subtle vocal inflection that suggests new distance in a previously close relationship.

Fever and cough indicate the presence of an infection or some other disadvantageous abnormal state. Anxiety and sadness arise from states that are disadvantageous, but generally not abnormal. This apparently major difference can be turned on its head by noting that the infections that arouse fever and cough are not exactly diseases, they are just conflicts with pathogens of the sort that our bodies manage constantly. The symptoms are aspects of the body’s well-established plan for dealing with infections. Both physical and emotional responses are useful only in certain situations. For physical responses these situations are more tangible and more likely to be abnormalities. For emotional responses, the etiology is not usually a disease process. To avoid confronting the complex social situations that arouse negative emotions, psychiatry has defined extremes of negative emotions as disorders. The result is a major emphasis on individual differences in “vulnerability” to negative emotions and a relative neglect of causes in the environment.

6.5 Evolution and emotions

The proper foundation for understanding emotional disorders is an evolutionary understanding of why the emotions exist at all (Nesse 1990; Tooby and Cosmides 1990; Nesse 1998; Nesse and Ellsworth 2009). The same logic is at the heart of pathophysiology. To understand the kidney, we first try to understand what it is for. Armed with this knowledge, we can understand how the nephron works and why it is the way it is. Such evolutionary functional understanding is so intrinsic to physiology that it is easy to overlook that it includes two separate kinds of knowledge, one an evolutionary explanation for why a trait exists at all, the other a proximate explanation for the details of the trait’s structure and how it works (Mayr 1961).

It is tempting to posit functions for emotions that are just as straightforward as functions for abdominal organs, but this is a mistake. The abdominal organs are always present and constantly useful, while emotional states are aroused only in certain situations and they are useful only in those situations. Panic, for instance, may be life-saving when serious danger is present, otherwise it is worse than useless. The correct way to analyze the utility of an emotional state
is to define the situations in which it is useful and the adaptive challenges posed by those situations. In the face of life-threatening danger, rapid breathing oxygenates the blood, muscle tension increases strength, and insulin allows glucose to flow into muscles. Emotions have utility for communication, motivation, and for adjusting physiology and behavior, but there is no need to consider which of these is primary. All are part of a special coordinated state that gives an advantage in a certain situation (Nesse 1990; Tooby and Cosmides 1990). For instance, sweating, rapid heartbeat, muscle tension, and a wish to escape are all useful when confronted by dangers that demand fight or flight, and they serve a variety of related functions. Emotions are like computer programs that adjust multiple aspects of the organism to cope with the exigencies of situations that have recurred over evolutionary time. Organisms with such abilities to adjust have an advantage over those that make no adjustments.

Emotions are positive or negative for the simple reason that special states are useful only in situations that pose opportunities or threats. Positive or negative subjective experience is but one aspect of an emotional state that includes changes in arousal, motivation, physiology, memory, and action tendencies (Plutchik 2003). Negative emotions are naturally associated with untoward situations, so it is easy to incorrectly conclude that they are “themselves problems. This “clinician’s illusion” is a serious impediment to understanding and treating emotional problems (Nesse 2005; Nesse and Ellsworth 2009).

It would be grand if all who treat emotional disorders could take several courses about emotions or at least read one good textbook, such as Plutchik’s (2003), but some of the debates in emotions research would likely be more distracting than illuminating (Ekman and Davidson 1994). For instance, arguments continue about whether emotions are best viewed as dimensions or as a few distinct basic kinds with combinations. An evolutionary approach offers a possible resolution by tracing the phylogeny of various emotions over evolutionary time as they have been gradually but only partially differentiated from one another in order to cope with diverse kinds of situations (Nesse 2004).

This view has profound implications for psychiatric diagnosis and the comorbidity of emotional disorders. For instance, instead of attempting to determine whether the various anxiety disorders are fundamentally the same or fundamentally different, it suggests that anxiety has been partially differentiated into subtypes shaped to cope with a variety of different kinds of dangers. We should, therefore, not expect to be able to differentiate subtypes of anxiety sharply; the boundaries between them are blurred (Marks and Nesse 1994). Similarly, the profound overlap between anxiety, sadness, low mood, and depression arises because they are responses to related kinds of danger. Anxiety is aroused by
situations that pose threats of possible future loss. Sadness is aroused by loss. Low mood is aroused by the expectation that one will be unable to reach an important goal. The decreased motivation encourages seeking another strategy or, if nothing works, disengaging from pursuit of the goal. If efforts persist nonetheless, ordinary low mood is likely to escalate to clinical depression.

There is no room here for a detailed consideration of the full spectrum of emotions, to say nothing of the extensive research and writing about them (Barlow 1991; Izard 1992; Oatley and Johnson-Laird 1995; Lewis and Haviland-Jones 2000; Fessler 2003; Fessler and Haley 2003). Instead, consider a list of some common situations and the emotions they arouse:

- opportunity $\rightarrow$ desire, excitement
- success $\rightarrow$ joy, happiness
- failure $\rightarrow$ disappointment
- threat of damage $\rightarrow$ fear
- threat of social loss $\rightarrow$ anxiety
- loss $\rightarrow$ sadness
- failure to make progress towards an important goal $\rightarrow$ low mood
- inability to get or protect an essential resource $\rightarrow$ despair
- betrayal $\rightarrow$ anger
- contamination $\rightarrow$ disgust.

The list could be greatly extended, but the relationship among different emotions becomes clearer if they are organized into groups that correspond to the two main classes of situations individuals need to cope with (Nesse 1990, 2004). The first is goal pursuit and the problem of what to do when, and with how much effort and persistence. Living is a sequence of episodes in which organisms attempt to reach goals and avoid losses. Table 6.1 summarizes the emotions that arise in the situations associated with goal pursuit. It presumes that a somewhat consistent set of brain mechanisms has regulated the pursuit

<table>
<thead>
<tr>
<th>Situation</th>
<th>Before</th>
<th>During</th>
<th>Obstacle</th>
<th>After success</th>
<th>After failure</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Opportunity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social</td>
<td>Excitement</td>
<td>Engagement</td>
<td>Frustration</td>
<td>Joy</td>
<td>Disappointment</td>
</tr>
<tr>
<td>Physical</td>
<td>Desire</td>
<td>Flow</td>
<td>Anger</td>
<td>Happiness</td>
<td>Pleasure</td>
</tr>
<tr>
<td></td>
<td>Interest</td>
<td></td>
<td>Despair</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Threat</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social</td>
<td>Anxiety</td>
<td>Confidence</td>
<td>Dread</td>
<td>Relief</td>
<td>Sadness</td>
</tr>
<tr>
<td>Physical</td>
<td>Fear</td>
<td>Coping</td>
<td>Despair</td>
<td></td>
<td>Pain</td>
</tr>
</tbody>
</table>
of diverse goals in different organisms over hundreds of millions of years. For any particular species, these global emotions gradually become somewhat specialized to cope with particular kinds of goals. For instance, when faced with the possibility of losing a mate most humans experience not just generic anxiety, but the complex emotion of jealousy. The regulation of these emotions is further specialized by life experience.

The other group contains emotions shaped to deal with the situations that repeatedly arise in managing social relationships. As most readers will know, evolutionists and economists often model the trading of favors as a prisoner’s dilemma in which the maximum net outcome emerges from repeated mutual cooperation, but on any given move, a player who defects gets a big gain at the expense of the other player (Axelrod and Hamilton 1981). We and others have argued that these situations are so ubiquitous that they have shaped specific emotions: trust and friendship after repeated successful exchanges, suspicion and anger before and after the other defects, and anxiety and guilt before and after the self defects (see Table 6.2) (Ketelaar and Clore 1997).

These tables are not intended to be exhaustive. For instance, surprise is a more general emotion aroused by situations that give rise to unexpected outcomes. Disgust probably evolved to protect us from contaminated materials, but it seems to have been co-opted for use in the mechanisms that keep us away from those who are judged morally unclean. All of the above emotions deserve extended explanations that are available elsewhere (Nesse 1990; Plutchik 2003). They are summarized briefly here as a prelude to addressing the question of emotional disorders.

### 6.6 Emotional disorders

An evolutionary perspective on emotions has several implications for a nosology of psychiatric disorders.

1. Emotional disorders should be recognized as distinct from other mental disorders. They are, like chronic pain, abnormalities of the regulation of useful responses and thus very different from disorders such as psychoses that are abnormal in any amount and any situation (Watson 2005).
In DSM-II they were better unified, but they have since each been pulled out as separate disorders.

2. Because emotions adjust the organism to cope with certain kinds of situations, the normality of an emotional state cannot be assessed without information about the situation (except for certain extreme emotional states that will be abnormal no matter what the situation).

3. The word “disorder” implies an abnormality of the mechanisms that regulate emotions, for instance panic in safe situations. Such abnormal expressions of emotions must be carefully distinguished from emotions that arise from normal mechanisms but nonetheless cause distress or impaired function, such as depressive symptoms arising from a fruitless job search (Wakefield 1992).

4. Two global classes of abnormalities are possible for each emotion:
   a. Too much: too quickly aroused, too intense, too long, or aroused by nonspecific cues.
   b. Too little: too slowly aroused, too mild, too short, or aroused only by excessively specific cues.

5. Emotions researchers now recognize that emotions arise not from directly apprehended cues, but from an appraisal of what the new information means for an individual’s ability to reach personal goals (Ellsworth 1991), a perspective that encourages attention to the life of the individual.

6. Negative emotions are just as useful as positive emotions. It is essential to avoid the clinician’s illusion that makes all negative emotions seem abnormal and all positive emotions seem normal. No one comes to the clinic complaining of too little anxiety or an inability to feel sad, but this is just an artifact of our limited imagination and the absence of a scientific foundation for diagnosis of emotional disorders. People with these disorders exist, they just are not complaining or coming for treatment. Instead, they show up in the emergency room or jail or unemployment lines.

7. The mechanisms that regulate expression of emotions are governed by the smoke detector principle: inexpensive defenses are often subject to false alarms that are perfectly normal (Nesse 2005).

8. What is useful for our genes is not necessarily useful for our selves. Much normal emotion, especially negative emotion, may not be worthwhile for individuals at all, but only for their genes, and sometimes only for their genes in kin.

9. It is also important to recognize that some emotions may have been shaped in the Paleolithic, which may render them useless or even harmful in the
I EVOLUTIONARY FOUNDATIONS FOR PSYCHIATRIC DIAGNOSIS

modern environment, even though they arise from normal mechanisms. For instance, expressing normal anger towards one’s boss is likely to be maladaptive in a modern bureaucracy.

10. The distinction between negative and positive emotions intersects the distinction between abnormalities of excess and deficit to define four broad classes of emotional disorders, two of which have been neglected because they do not lead to subjective complaints. See Table 6.3 for details.

These and related principles provide a foundation for a scientific nosology for emotional disorders. An improved diagnostic system based on them will seem senseless to those who do not understand the behavioral biology of emotions.

This framework encourages systematic consideration of disorders of excess and deficiency for every emotion, not just anxiety and depression. The vast majority of treatment is for anxiety and depression, of course. They are usually called affects instead of emotions, to reflect their more enduring presence and the difficulty of connecting them to a very specific situation, but the conclusions are the same nonetheless.

An emphasis on the evolved utility of negative emotions should not lead to the conclusion that they are always useful, nor should it distract attention from the huge genetic variation in emotional predispositions. Some people rarely experience guilt while others feel constantly that they have somehow transgressed. Some people rarely worry, others worry constantly. Some people have never experienced romantic love, others fall madly in love with remarkable regularity. This variation poses a major problem for any attempt to determine what emotional experiences are normal.

Part of the answer is in how natural selection shapes the systems that regulate behavior. About half of the variation among individuals in most emotional traits arises from genetic differences. Why hasn’t natural selection shaped a much more narrow range of responsiveness that we can recognize as “normal”? It is because humans have evolved in varying physical and social environments, so variations for a substantial range around the mean may not have a consistent effect on fitness. The resulting variation in personality traits is so large as to sometimes make us wonder if we are even justified in talking about human nature (see also Chapter 7).

Table 6.3 Categories of emotional disorders

<table>
<thead>
<tr>
<th></th>
<th>Excess</th>
<th>Deficit</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Positive emotions</strong></td>
<td>Mania, erotomania</td>
<td>Lack of joy, love, interest</td>
</tr>
<tr>
<td><strong>Negative emotions</strong></td>
<td>The usual emotional disorders: anxiety, depression, etc.</td>
<td>Deficits of anxiety, low mood, jealousy, etc.</td>
</tr>
</tbody>
</table>
We can now return to the DSM approach to diagnosis and the problem of taking context into account. The criteria for some disorders have built-in exclusions that generally ensure that anyone who meets criteria does indeed have a disorder. For instance, the criteria for panic disorder refer to “unexpected attacks”, which excludes panic in life-threatening situations. Panic disorder is a reliably pathological condition in which the threshold for panic is so low that attacks emerge spontaneously. What an evolutionary perspective adds is recognition that panic is a normal response that is expressed too readily in panic disorder (Nesse 1987). This simple fact is useful in psychotherapy. Patients who have spent months fearing they have heart disease or a brain tumor often can be helped to recognize that their symptoms would indeed be useful in extreme danger and that they are experiencing mere false alarms.

Jealousy is a more complicated example. In the face of threats to a mate’s fidelity, jealousy is normal and its absence is abnormal (Buss 2000). However, in many instances jealousy seems to be pathological. In many such cases, later evidence reveals the emotional response was an accurate indication of what was actually happening. In others, jealousy is aroused in someone who is depressed or who otherwise feels that his or her partner could do better with someone else. Then there is the psychoanalytic observation that jealousy can arise from projecting illicit desires onto an innocent partner. The important point here is that different cases of pathological jealousy may have different origins, but differentiating and understanding them requires knowing the situations in which the emotion is useful.

The overwhelmingly common disorder is, of course, depression. Increasingly, patients receive a diagnosis after a brief interview with a general physician, who prescribes antidepressants and advises a return visit in a month. Such perfunctory treatment is often justified by noting that the patient has met criteria for a pathological condition, major depression, whose presumed etiology is a deficiency of brain neurotransmitters. Drug treatment seems indicated and has been proven somewhat effective, so why not get on with it? This sequence completely ignores any possible utility of low mood, to say nothing of the causes of an individual’s depression.

An evolutionary approach recognizes that low mood is useful to disengage effort from enterprises that are failing (Price and Sloman 1987; McGuire et al. 1997; Nesse 2000; Wrosch et al. 2003; Nettie 2004; Nesse 2009). If the person persists in useless efforts, the low mood escalates to full depression. It sounds easy to recommend giving up a fruitless pursuit until you realize that the goal may be getting a child off drugs, finding a job, or ending an affair. Treating depression without a careful examination of a patient’s motivational structure is like treating a cough without first trying to find its cause (Nesse 2005, 2009).
6.7 The importance of analyzing motivational structure

The most useful contribution evolution makes to classifying, diagnosing, and treating emotional disorders may be the framework it offers for analyzing the motivational structure of an individual’s life. Emotions arise from perceived problems and opportunities in the motivational structure. Like other organisms, humans must allocate three kinds of effort to get resources in six different areas. Somatic effort yields personal resources and material resources. Reproductive effort yields mates and offspring. Social effort yields allies and status. Many people seem to imagine that there is some normal way to live without compromises, but an evolutionary perspective reminds us all that every human action is an investment in getting one kind of resource at the expense of others. More time working out means less time working. More time impressing potential mates means less time for childcare. More time seeking status means less time for everything else.

The motivation regulation system seems to be designed, sensibly enough, to focus effort where it is most needed, that is, wherever it will yield the greatest payoffs of reproduction-limiting resources for the least investment. It would be so nice if our minds settled comfortably to a focus on what we have, but after any satisfaction the mind turns quickly to solving the next problem (Nesse 2004). Many tasks are simply enterprises that work well, such as a job or a marriage. What then is a life problem? A life problem is a difficulty in getting or keeping some important resource. People describe their problems in such diverse ways it is at first amazing to see how easily they all fit into a behavioral biological framework. The foundation for any therapy, especially psychotherapy, is a detailed understanding of what resources and sources of resources the person has, what he or she wants, how he or she is going about reaching these goals, and what the expectations are for success or failure. Many depressed people seem to have nothing major lacking in their lives, but as we get to know them, we find that they are striving to get love from a cold mother, sex from an uninterested spouse, or praise from a competitive boss. Or, they are trying to be truly good at all times, or to be the world’s best in some status competition, achievements that are rare, and always temporary.

Good clinicians intuitively grasp motivational structures and the exigencies that give rise to an individual’s problems. An evolutionary perspective and knowledge about emotions can help nearly every therapist to do this even better. There is the risk, of course, of using such insights to make crude suggestions. A patient who visited the emergency room attributed his depression to his wife’s disinterest in sex. He was told, “Well, you will have to leave her or put up with it, those are your choices.” Better therapists know that people have good reasons for why they live in the way they do. They examine their patients to see if symptoms arise from bipolar disorder or some other distinctive
condition, but they recognize that diagnoses are no substitute for a deep understanding of a person's life.

6.8 Towards an evolutionary foundation for psychiatric nosology

The crucial missing ingredient for a truly medical nosology for emotional disorders is a functional understanding of the emotions and their regulation that is comparable to the functional understanding that physiology provides for the rest of medicine. Brain mechanisms are an essential part of this missing knowledge, but they are no more complete in themselves than the anatomy and mechanisms of the kidneys are for understanding the causes of renal pathology. Understanding the adaptive utility of a system is just as important for emotional as for physiological systems. Evolution provides the missing functional perspective for understanding the emotions and their disorders.

Many readers may agree with much of the above argument and yet find themselves asking, “Yes, the problems are large and clear, but how can we craft a DSM-V that avoids them?” A straightforward approach is to classify emotional disorders in the same way medical disorders are classified, based on the etiological factors that give rise to them. As already noted, medical symptoms are usually aroused by fairly specific tissue-changing pathologies, while emotional symptoms are most often aroused by untoward social situations that are much less susceptible to neat classification. I think it is likely that finding reliable and valid categories for emotional disorders has been difficult because they are not distinct diseases with specific causes. They arise from interactions between neural and cognitive diatheses interacting with inherently subjective appraisals of complex situations.

Sometimes, as in bipolar disorder or panic disorder, nearly all the variance is in genetic individual differences. In most cases, however, the circumstances giving rise to the emotion also play a major role. Although such situations are diverse, they can be categorized nearly as neatly as the aversive emotions. Here are a few examples of some of the causal situations that clinicians recognize intuitively:

- unrequited love (inability to give up a hopeless romantic goal)
- unable to find an intimate partner
- unable to leave an unsatisfactory intimate relationship
- unable to find a job anywhere near as high status as one's parents
- unable to leave an unsatisfactory job
- personality disorder that disrupts adaptation in multiple domains
- being blackmailed
• unable to help a child in trouble
• health problems that prevent functioning in crucial roles
• an affair that threatens major relationships
• partner may be having an affair
• partner is ill or disabled.

If someone is shivering, we do not look to the brain center that mediates shivering for an explanation, we instead look at the temperature, clothing, possible infection, etc. There is variation, both innate and acquired, in how readily different people shiver, but this is only part of the picture. We don’t know what proportion of patients in our clinics have disorders of emotion regulation, and what proportion have basically normal mechanisms interacting with untoward circumstances. We need to know. Axis IV calls attention to life events. But because these events are carved out from consideration in reaching an Axis I diagnosis, diagnosis in psychiatry remains fundamentally different from that in the rest of medicine. General physicians no longer diagnose “cough disorder”, they use different diagnoses depending on the etiology. DSM-V should incorporate life events and life situations into the main diagnostic categories, where their role as elicitors of emotions will be clearer.

Even stating the problem as distinguishing between individual differences or environmental effects is a mistake. Every emotional disorder arises from interactions among an individual’s brain mechanisms, cognitive patterns, and his or her appraisals of the significance of information for reaching personal goals. The first variable is influenced by genes, early experiences, drugs, and other direct influences on brain mechanisms. The cognitive appraisal is influenced by personal and cultural experiences as well as individual idiosyncrasies from many sources. The events that arouse emotions arise from complex socio-cultural contexts, but also from the social network that grows around an individual, which is influenced by all the other variables. This is complicated. There is no getting around the complexity without excluding important factors or causal links.

We began with the supposition that continuing consternation about psychiatric nosology suggests that we are missing something basic. We are trying to categorize emotional disorders without a foundation of the understanding of the emotions and their origins and functions. This foundation illuminates many of the problems encountered by nosologies for emotional disorders. Unfortunately, however, it does not offer a simple solution. Instead, it shows that extreme emotional states arise not from one source, but from interactions of individual brain differences with complex life circumstances interpreted by diverse cognitive appraisals and psychological defenses. The categories of cleanly differentiated well-defined emotional disorders that we have been seeking do not exist (Nesse and Ellsworth 2009).
One could conclude from this that the DSM-IV approach to emotional disorders is about the best that can be done. This would be like nineteenth century physicians being satisfied with the diagnosis of “fever” because they can measure it reliably even though they don’t know its causes. Instead, we need to proceed in the same way general physicians approach symptoms. They consider all possible causes in a differential diagnosis, then they investigate to find the etiology in any particular case. For emotional disorders, we must investigate the motivational structures of individuals in the same kind of detail that has been lavished on brain mechanisms.

To escape from abstractions, consider three cases.

---

**Case 1**
This 35-year-old woman has moderate depression and anxiety with intense anger and jealousy.

**Situation:** She learned her spouse is having an affair and wants to leave him, but she has no income and would have to give up her friends and her art career.

**Person and vulnerability factors:** Somewhat emotional in general, she has a slight tendency towards negative affect, but no enduring abnormal regulation of emotions in general and no family history of mental disorders.

**Etiology:** Her emotions are normal responses to her life situation.

---

**Case 2**
This 35-year-old woman has moderate depression and anxiety with intense anger and jealousy.

**Situation:** She suspects her spouse is having an affair but has no evidence of this. He denies it and tries to reassure her.

**Person and vulnerability factors:** She has always believed men will prefer other women and has been pathologically jealous in most of her relationships. She attributes this to her father leaving her mother when she was 5 years old. No family history of emotional disorder.

**Etiology:** Personality problem likely related to early life events; intense jealousy and other emotions are secondary.

---

**Case 3**
This 35-year-old woman has moderate depression and anxiety with intense anger and jealousy.

**Situation:** She accuses her spouse of having affairs, but only after she has been without sleep for several days, often while drinking.

**Person and vulnerability factors:** Strong family history of bipolar disorder.

**Etiology:** Genes causing bipolar disorder, complicated by alcoholism, relationship problems, and extreme jealousy.
These cases illustrate what most clinicians know: the same clinical conditions can arise from fundamentally different causes. It therefore makes no sense to view these emotions as a specific disorder. In case 1 the symptoms arise from an untoward situation, in case 2 from a personality disorder, in case 3 from bipolar disorder. Every clinician will think of more realistic and complex cases, for example a man with bipolar tendencies, chronic relationship difficulties, a low threshold for jealousy and anger, who drinks heavily and is having an affair.

The implications for the DSM-V are substantial, but not simple. Detailed consideration of the opportunity by the DSM-V Committees is indicated. It is essential to recognize that an evolutionary foundation is fully compatible with other biological and medical approaches. The DSM-IV has encouraged much useful work on the problem of why some people have tendencies to excessive anxiety and depression, and the brain mechanisms that mediate affects. Evolution puts this knowledge in perspective by emphasizing that these affects can be normal, their regulation mechanisms were shaped by natural selection, and there are likely good evolutionary reasons why these mechanisms are so vulnerable to failure. It also highlights the need to look for disorders of regulation for all emotions, especially the neglected disorders characterized by deficient negative or excessive positive affect.

Some people think that the utility of negative emotions means that they should not be treated. This is a serious mistake. We have much to learn from general medicine, where both the utility and the harm caused by responses such as pain and diarrhea is well recognized, and where relief of suffering by blocking defensive responses is a routine goal of clinical work, whether the symptom is being aroused normally or arises from a faulty mechanism. Campaigns to convince the public and practitioners that depression and anxiety are brain diseases have motivated much useful research and have decreased stigma, but they are biologically naïve. An evolutionary approach supports a more medical model in which clinicians recognize many symptoms as defenses shaped by natural selection that are aroused by more primary causes, and others arising from defects in the systems that regulate defenses. The clinician tries to identify and remove the factors arousing the symptoms when possible. When that is not possible, a good psychiatrist tries to relieve suffering, often by using drugs to block normal responses. If that is not possible, then the clinician tries to relieve suffering, even if that means using drugs to block normal defensive responses. Evolutionary biology offers a biological foundation for a genuinely medical model for understanding and diagnosing emotional disorders.

References
American Psychiatric Association (1952) Diagnostic and Statistical Manual of Mental Disorders. American Psychiatric Association, Washington, DC.


REFERENCES


