Measurement for a Human Science*

JOHN MIROWSKY
CATHERINE E. ROSS
Ohio State University


This paper argues a number of points about measurement in the sociology of mental health: (1) measurement is critical, (2) measures should represent and assess elements of human experience, taking measure of life as people feel it, sense it, and understand it, and (3) social scientists should create a human science, producing information for the people it studies so that they can better understand and control their own lives. We argue that a human science is best achieved with the use of indexes, not diagnoses, to measure mental health. We present a brief history of diagnostic instruments and detail how a diagnosis is made. We show how use of diagnoses to measure mental health discounts much human suffering. They dichotomize the true range of feelings and emotions into crude either/or distinctions that do not reflect the reality of people’s lives, and they often exclude suffering such as that due to loss or illness that does not meet medical model preconceptions about mental disorder. Using diagnoses to measure mental health presents a reified image of hidden disease knowable and manageable only by trained professionals—beyond the capacity of the suffering individuals to understand and control.

Measures of mental health should represent and assess elements of human experience clarified and refined from that experience but not removed from it. Research in mental health speaks most true when it takes measure of life as people feel it, sense it, and experience it.

MEASUREMENT AS CRITICAL

How can we know what a person is feeling or thinking? Psychological distress is probably one of the most difficult concepts to measure. Difficulty in measurement means that researchers should put a great deal of thought and effort into measurement, not abandon measurement. Unless we measure levels of depression, anxiety, anger, and so on, we can only speculate about their associations with social conditions. Does divorce free women from the restrictions and limitations of traditional power relationships in which men have autonomy and women are dependent and oppressed, so that divorce increases hope, happiness, and joy? Or, on the contrary, does divorce increase women’s economic hardship and child care burdens and decrease security and support, thereby leading to distress, demoralization, and despair? In order to answer questions like these, researchers need to measure concepts of distress, demoralization, hope, and joy. Without measurement, our answers will simply reflect our preconceptions, biases, and ideology.

Measurement is our one hope of escaping preconception. Without it we cannot distinguish shared observations from shared beliefs. To some social scientists this makes measurement an onerous limitation on the collective definition of truth. To others it liberates their science to discover in observations phenomena...
that were unexpected, unimagined, unknown, and in some cases perhaps even unwelcome. Measurement keeps us honest, with ourselves and with each other. As runners, the stopwatch has forced an unwelcome truth on us more than once. The same can be said of the bathroom scale. We human beings often project our hopes, fears, and prejudices into our perceptions. Reference to an external standard helps minimize that bias.

Measurement is the link between concepts and reality. The act of gauging attributes by reference to standards brings the abstract realm of ideas into contact with the concrete and practical realm of being. Measurement is critical. When we choose to measure, we choose to see that which we may not expect, prefer, or understand. When we choose what to measure, we choose the medium of our common view.

Our measures of mental health can take one of two forms. In the first, measures represent the reality of people’s lives, feelings, emotions, and thoughts as perceived and reported by the people themselves. Here, indexes assess the type of problem (like depression, or anxiety, or anger), and the degree of severity or frequency of the problem along a continuum, as reported by the person. In the second form, measurement represents the reality of psychiatric preconception that mental illness is an unseen biological condition of the body that is either present or absent, ignoring the subtlety of feelings that people experience. Here, psychiatric diagnoses categorize people into, for instance, depressed or not depressed, anxious or not anxious, according to diagnostic decision rules that represent psychiatric preconceptions rather than people’s own feelings and cognitions. In so doing, diagnoses build in medical model preconceptions that social conditions such as inequality, poverty, or hardship are not the causes of psychological problems. As we will argue, indexes have advantages over diagnoses in the assessment of psychological problems.

SOCIOMETRY OF MENTAL HEALTH AS A HUMAN SCIENCE

The human beings we, as sociologists, study are different from every other object of study. Humans communicate, think, decide, plan, and act. Knowledge gives people the tools to act effectively to improve their lives. Scientists help create that knowledge. Physicists who study electromagnetic energy develop the ability to control it, and they market that ability to government and industry. Physicists may have scruples about military applications and environmental impacts, but they need not worry about infringing on electromagnetism’s self-determination. It has none. In sociology things are different. Sociologists can create knowledge for professionals or create knowledge for the people we study. As scientists, we will create knowledge that is power for someone. Who do we see as the consumers of our product? For whom do we create knowledge?

Many researchers want to create knowledge that helps ordinary humans understand and control their own lives, so that those ordinary humans can avoid suffering, avoid making others suffer, and achieve the well-being they desire. That goal influences a researcher’s logical choice of scientific questions and methods. Measurement for a human science has the goal of producing information that people can use themselves; not information that professionals, like psychiatrists, can use to manipulate emotions, cognitions, and behaviors of others, often with drugs. We do not see our job as market research for psychiatric services; we see it as providing information that might be useful to the people themselves.

Placing people in diagnostic categories treats them as objects and ignores the persons, or treats the persons as mere epiphenomena of the biological mechanisms of their thoughts, feelings, and actions. Operating premises such as those make sense if you want to develop a science that specialists use to alter the moods, thoughts, and behaviors of others. The goal of placing people in diagnostic categories is often to show an unmet need for psychiatric services. Such research estimates the number of people who “need” psychiatric services because they qualify for a diagnosis but are not getting services (Kessler and Zhao 1999). placement of psychiatric disorders among the general population is flawed and misleading because services are not used. Current psychiatric epidemiology is flawed and misleading because they qualify for a diagnosis but are not getting services (Kessler and Zhao 1999).
with psychological problems who are not receiving psychiatric treatment do not want to seek help, do not feel comfortable consulting a psychiatrist, and would rather deal with their problems on their own (Lin and Parikh 1999). "Unmet need" is defined by the psychiatrists, not the people themselves.

Psychiatric diagnosis of psychological problems furthers the goals of putting control in the hands of professionals. In contrast, a human science creates information for the people it studies so that they can better understand and control their own lives.

THE LANGUAGE OF MEASUREMENT AS CLEAR AND DIRECT, NOT REIFIED AND OBSCURE

A human science informs the people whose lives it studies. To communicate well we need to talk about things many persons experience using terms that most persons can understand. Arcane concepts and obscure terms lend an aura of authority. They establish dominance and subordination. They elevate the professional by implying, "This is over your head. Put yourself in my hands."

The following examples from internal medicine illustrate that researchers can create information either for the general population or for their doctors. Both kinds of information can be effective, but only one kind gives control over the problem to the individual who is suffering. Imagine this scene. An older man tells the doctor that his leg is giving him problems; it sometimes gets lame and is painful when he walks. The doctor tells the man he has intermittent claudication. The man thinks that it sounds serious, and that it is a good thing he came to the doctor. However, intermittent claudication is merely the diagnostic label for occasional lameness along with pains in the leg when walking. By stating the medical diagnosis the doctor merely repeats what the patient said, but in the terms used in medicine. The symptoms do suggest a serious problem. In most cases it results when a fatty plaque in an artery cuts off the supply of blood to part of the leg, starving and perhaps killing some of the muscle. It also suggests that similar plaque may be forming in the arteries that feed the heart muscle or the brain. Individuals often cut back on walking when they start getting the pains and lameness, but they should do just the opposite: work up to walking an hour every day. In most cases that will eliminate the leg pains, generally improve circulation throughout the body, and probably also slow the growth of fatty plaques. In most cases the problem could be avoided completely with a lifestyle that includes regular exercise and a reasonable diet.

Proponents of medical explanations and drug treatments for physical and psychological problems often seem blind to the idea of effective information that can be used by the persons themselves to avoid or solve their own problems. Levenstein (1998, 2000) presents another example from internal medicine. The association of heliobacter pylori with peptic ulcer leads to a disregard of ulcers' association with the stress or hard labor of lower status jobs, and the effectiveness of antibiotic treatment often obviates the search for things people can do themselves such as eating breakfast, getting enough sleep, quitting smoking, or avoiding heavy drinking. For proponents of the medical model in psychiatry, the effectiveness of selective serotonin reuptake inhibitors for emotional problems labeled dysthymia invalidates and obviates the search for ways that people themselves can prevent, correct, or manage such problems. Just because drugs work does not mean that biological explanations are the only valid or effective understanding of the problem. Viewing psychological problems as diseases does, however, give control over these problems to the medical and drug industry, rather than to the people themselves. In contrast, a human science creates information for the people it studies so that they can better understand and control their own lives.

How should a human science measure mental health?

THE ADVANTAGES OF INDEXES OVER DIAGNOSTIC CATEGORIES

Real but not Dichotomous

Psychological problems are real, but they are not discrete. They are not something that is entirely present or entirely absent, without shades in between. Psychological problems are not entities. They are not alien things that get into a person and wreak havoc. Nevertheless, psychiatrists speak of depression and other
psychological problems as if discrete entities enter the bodies and minds of hapless victims. The psychiatrist detects the presence of an entity and determines its species (makes a diagnosis), then selects an appropriate weapon against it (usually drug treatment). The imagery of detection follows from the language of discrete entities. This categorical language is the legacy of 19th century epidemiology and microbiology. A person is diseased or not. The disease is malaria or not, cholera or not, schistosomiasis or not. A language of categories fits some realities better than others. It fits the reality of psychological problems poorly (Mirowsky and Ross 1989a, 1989b).

Assessing the Type and Severity of Problems

Instead of diagnosing people, we can assess the type and severity of symptoms using indexes or scales. The fact is that we do not have to place people in diagnostic categories in order to know which subpopulations suffer more than others. Counting the number of persons in a diagnostic category is easily replaced by counting the number of symptoms of a particular type that various people have. The later strategy avoids the proliferation of diseases, each with its own name and mythical status as a unique, discrete entity. We need to remember, though, that a category of symptoms is a mental pigeon hole too. People are the real entities. The symptoms are merely things that some people feel or think or do more than others, for reasons we would like to know. Some of those things appear together more frequently than others, and those are the ones we treat as a single type of symptom.

It is useful to think in terms of the type and severity of psychological problems (Mirowsky and Ross 1989a, 1989b). Depression is a type of psychological problem. So is anxiety. Each type of problem ranges on a continuum from not at all severe to very severe. People score at all points on the continuum—from very few symptoms to many symptoms. People can get a severity score for each type of psychological problem. Contrast this with the diagnostic approach. Imagine two people on either side of some arbitrary cut-off that defines depression. One has just enough symptoms to get a diagnosis, and the other is just short of enough. Although the type and severity of their problems are very similar, one is diagnosed as depressed and the other is not. The diagnostic imposition ignores their similarity. Imagine another two people. One is happy, fulfilled, and productive. The other is demoralized, hopeless, and miserable, but just short of meeting the criteria for a diagnosis of depression. The diagnostic imposition ignores their differences. Diagnosis throws away information on the similarity of some cases and on the dissimilarity of others (Mirowsky and Ross 1989a, 1989b).

Unreliability of Dichotomous Measures

Throwing away information doesn’t help us understand problems, it hinders us. Many people erroneously believe that the accuracy of an assessment is improved by making crude distinctions. For example, a typical bathroom scale measures weight accurately within a range of plus or minus 2 pounds. This means that if a person’s true weight is 140 pounds, there’s a 99 percent probability that the bathroom scale will say the person’s weight is in the range from 138 to 142 pounds. It is unlikely the scale will give the person’s exact, true weight, and it would be sheer chance if the scale did. Nevertheless, the bathroom scale is almost certain to register within 2 pounds of the person’s true weight. Although the bathroom scale will rarely tell us a person’s true weight, it will usually tell us correctly which 5-pound range the person’s weight is in. The broader the ranges, the more likely a bathroom scale tells us correctly in which range the person’s weight is. The broadest range would divide everybody into two categories: heavies and not heavies. Almost everyone would be correctly classified. The current bathroom scale, which almost never shows a person’s true weight, could be improved so that it only shows whether a person is heavy or not. Now suppose you are trying to lose weight. Which bathroom scale would you choose? The old one that is almost never correct, or the new one that is almost always correct?

The fallacy of the two-category scale lies in confusing certainty and reliability (Mirowsky and Ross 1989a, 1989b). In psychometrics (the science of measuring psychological states and traits), reliability is the exactness of reproduction that can be achieved with a given measure. Reliability in psychometrics is analogous to fidelity in electronics. Fidelity is the degree to
which a system, such as a radio or television, accurately reproduces the essential characteristics of its input. Reliability increases with the precision of assessment. A measure of length is more reliable if the ruler is marked in inches than if the ruler is marked only in feet. Reliability is lowest if measurement is dichotomous—a simple yes or no, in or out, heavy or not, diagnosed or not (Cohen 1983).

As an assessment becomes broader, it becomes less sensitive to meaningful changes or differences, and the ratio of information to noise declines. This may seem strange, given that broader categories increase the certainty of an assessment. Once again, though, it is reliability and not certainty that we need. After all, if we reduced everything to a single category, our certainty would be perfect but meaningless. The bathroom scale that has a random error of plus or minus two pounds has exactly the same amount of random error when we paint out all the little marks and replace them with a red zone for heavy and a white zone for not heavy. The crude split eliminates almost all of the information without eliminating any of the random error.

Imagine being asked, “Are you depressed—yes or no?” You think about how to answer. You feel depressed at times and are not enjoying life very much, you don’t feel too hopeful that things will get better, and sometimes you lie awake at night, troubled; on the other hand, you get out of bed every day, go to work and care for your children, you don’t feel unable to concentrate at all, and you don’t feel totally alone or think about death. Are you depressed? The crude dichotomy does not allow you to give a meaningful answer about how you feel. A more accurate assessment would allow you to tell the interviewer about the frequency of each symptom more exactly. Only people in perfect mental health, on the one hand, or extremely depressed to the point of considering suicide, on the other, could easily answer yes or no. For the large majority, the dichotomy is too insensitive to describe their subjective experience. Yet diagnostic instruments often force people to answer yes or no to questions about how they feel (See Table 1).

**Insensitivity of Diagnoses**

When the full range of symptoms is split into two categories (enough for a diagnosis of depression, or not enough), most of the information is lost, but all the random error remains. A diagnosis of depression is a profoundly insensitive measure. As a consequence, it can be difficult to find meaningful changes or differences in diagnosed depression. For example, a community study finds education and family income do not predict whether a person gets a diagnosis of depression or not (Weissman 1987). One of the researchers concludes that “depression equally affects the educated and uneducated, the rich and poor, white and black Americans, blue and white collar workers” (Weissman 1987:448). Nothing could be further from the truth. No theory, whether social, psychological, genetic, or environmental, predicts that the poor, the uneducated, the blacks, and the blue collar workers have the same exposure to the causes of depression as the rich, the educated, the whites, and the white collar workers. With a sufficiently insensitive measure, we cannot hear the suffering of millions, and cannot see the causes.

Researchers who use diagnostic categories to assess depression usually find that diagnostic categories are harder to predict than indexes. For instance, Turner and Lloyd (1999) predict depression two ways, with an index of depressive symptoms (CES-D) and with a diagnosis of major depression using diagnostic criteria set out in the Diagnostic and Statistical Manual of Mental Disorders (DSM III-R) of the American Psychiatric Association (1987). The patterns are similar, but they are weaker in the prediction of diagnosis. For instance, using a continuum of depressive symptoms, they find that women feel more depressed than men, persons of low socioeconomic status feel more depressed than those with higher status, people who are married feel less depressed than the divorced and widowed, and people with high levels of mastery and self-esteem feel less depressed than those with lower levels of mastery and self-esteem. In the prediction of a diagnosis of major depression, all the patterns are pretty much the same, but less is explained, and the effects are smaller and less significant. This is because the diagnosis is a crude measure. Most of the real information is lost or never assessed. The reliability of the measure is much lower. Yet Turner and Lloyd (1999) conclude, “Substantial similarities across these two sets of results suggest that the stress process is relevant to the occurrence of
TABLE 1. The Diagnostic Interview Schedule: Major Depression

1. “In your lifetime, have you ever had two weeks or more during which you felt sad, blue, depressed or when you lost all interest and pleasure in things that you usually cared about or enjoyed?”
   Yes____ No____

2. “Has there ever been a period of two weeks or longer when you lost your appetite?”
   Yes____ No____

3. “Have you ever lost weight without trying to—as much as two pounds a week for several weeks (or as much as ten pounds altogether)?”
   Yes____ No____

4. “Have you ever had a period when your eating increased so much that you gained as much as two pounds a week for several weeks (or ten pounds altogether)?”
   Yes____ No____

APPETITE SUMMARY: CHECK YES____ IF YES IN 2, 3, OR 4.

5. “Have you ever had a period of two weeks or more when you had trouble falling asleep, staying asleep or with waking up too early?”
   Yes____ No____

SLEEP SUMMARY: CHECK YES____ IF YES IN 5 OR 6.

6. “Have you ever had a period of two weeks or longer when you were sleeping too much?”
   Yes____ No____

TIRED OUT SUMMARY: CHECK YES____ IF YES IN 7.

7. “Has there ever been a period lasting two weeks or more when you felt tired out all the time?”
   Yes____ No____

8. “Has there ever been a period of two weeks or more when you talked or moved more slowly than is normal for you?”
   Yes____ No____

9. “Has there ever been a period of two weeks or more when you had to be moving all the time—that is, you couldn’t sit still and paced up and down?”
   Yes____ No____

SLOW, RESTLESS SUMMARY: CHECK YES____ IF YES IN 8 OR 9.

10. “Was there ever a period of several weeks when your interest in sex was a lot less than usual?”
    Yes____ No____

SEX SUMMARY: CHECK YES____ IF YES IN 10.

11. “Has there ever been a period of two weeks or more when you felt worthless, sinful or guilty?”
    Yes____ No____

WORTHLESS SUMMARY: CHECK YES____ IF YES IN 11.

12. “Has there ever been a period of two weeks or more when you had a lot more trouble concentrating than is normal for you?”
    Yes____ No____

13. “Have you ever had a period of two weeks or more when your thoughts came much slower than usual or seemed mixed up?”
    Yes____ No____

TROUBLE THINKING SUMMARY: CHECK YES____ IF YES IN 12 OR 13.

14. “Has there ever been a period of two weeks or more when you thought a lot about death—either your own, someone else’s or death in general?”
    Yes____ No____

15. “Has there ever been a period of two weeks or more when you felt like you wanted to die?”
    Yes____ No____
TABLE 1. (Continued)

16. "Have you ever felt so low you thought of committing suicide?"
   
   Yes____  No____

17. "Have you ever attempted suicide?"
   
   Yes____  No____

DEATH SUMMARY: CHECK YES____ IF YES IN 14, 15, 16, OR 17.

CHECK ITEM. REFER TO SUMMARY BOXES.

   _____ FEWER THAN 4 SUMMARY BOXES CHECKED YES. [No diagnosis of major depression.]
   _____ 4 OR MORE SUMMARY BOXES CHECKED YES, AND NO IN QUESTION 1. Go to question 20.
   _____ 4 OR MORE SUMMARY BOXES CHECKED YES, AND YES IN QUESTION 1. Go to question 18.

18. "You’ve said you’ve had a period of feeling depressed (blue, sad, etc.) and also said you’ve had some other problems (MENTION ALL CHECKED YES IN 2–17). Has there ever been a time when the feelings of depression and some of these other problems occurred together, that is within the same month?"
   
   Yes____  No____

19. "[If “no” to question 18]: So there’s never been a period when you felt depressed at the same time you were having some of these other problems?"
   
   Yes, has been a period____
   No, never been a period____ [No diagnosis]

20. "You said you have had periods when (MENTION ALL CHECKED YES IN 2–17). Was there ever a time when several of these problems occurred together—that is, within the same month?"
   
   Yes____
   No____ [No diagnosis]

21. "[If “yes” to question 20]: When you were having some of these problems at about the same time, were you feeling okay, or were you feeling low, gloomy, blue or uninterested in everything?"
   
   Low or equivalent____
   Okay____ [No diagnosis]

22. "[If “low or equivalent” to question 21]: What’s the longest spell you’ve ever had when you felt depressed (blue, sad, etc.) and had several of these other problems at the same time?"
   
   More than 2 weeks____
   Less than 2 weeks____ [No diagnosis]

23. "Have you had more than one spell when you felt depressed (blue, sad, etc.) and had several of these other problems at the same time?"
   
   Yes____  No____

24. "Did this spell (or any of those spells) occur just after someone close to you died?"
   
   Yes____  No____

25. "Have you had any spell of depression along with these other problems (MENTION SOME PROBLEMS CHECKED YES IN 2–17) at times when it wasn’t due to a death?"
   
   Yes____
   No____ [No diagnosis]

26. "Are you in one of these spells of feeling low or disinterested and having some of these other problems now?"
   
   Yes____  No____

27. "When did your last spell like that end?"
   
   Within last 2 weeks____
   Within last month____
   Within last 6 months____
   Within last year____
TABLE 1. (Continued)

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Now I’d like to know about the time when you were feeling depressed (sad, blue, etc.) for at least 2 weeks and had the largest number of these problems at the same time. How old were you at that time?”</td>
<td></td>
<td></td>
</tr>
<tr>
<td>____ years old</td>
<td></td>
<td></td>
</tr>
<tr>
<td>“During that spell of depression (blues, sadness, etc.), which of these other problems did you have? For instance, during that spell:”</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29. “Did you lose your appetite?”</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>30. “Did you lose weight without trying to—as much as two pounds a week for several weeks (or as much as ten pounds altogether)?”</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

(Questions 29–44 are the same questions as 2–17, phrased in the format shown in 29–30. Respondents are only asked about these symptoms during the worst spell if they answered “yes” to them in questions 2–17; that is, if they answered earlier that they had ever had them for a period of more than two weeks).

(Questions 24 and 25 show the exclusion of depression due to death of a loved one from the diagnosis of depression. Other exclusion criteria are not shown, but are assessed by probes in which the interviewer questions whether symptoms are due to physical illness, drugs, or other psychiatric problems).

Depressive disorder as well as for depressive symptoms . . . these results support the argument that risk and protective factors of relevance for depressive symptoms are likely to also be relevant for depressive disorder. On the other hand, the differences between these two sets of results are consistent with expectation based on the assumption that biological and genetic factors are also implicated in the occurrence of depressive disorder” (p. 391). In contrast to their interpretation, we argue that depressive symptoms are not something different than depressive disorder, but that reification of the diagnosis of depression obscures this fact. Depressive symptoms and depressive disorder are assessed using the same symptoms—feeling depressed, sad, a loss of interest in things, sleep problems, fatigue, feelings of worthlessness, trouble concentrating, and so on. The difference is that indexes assess the frequency or intensity of each symptom and don’t use dichotomous cut-offs and arbitrary duration rules (such as the symptom must be present for more than two weeks). Inability to predict the diagnostic category from social stressors is probably due to unreliability in the measure (Cohen 1983), not that depressive disorder is something different from severe symptoms of depression. Reification of diagnoses obscures the fact that the symptoms on which a diagnosis are based are the same as the symptoms used in indexes. Diagnoses are measured by researchers, just like any other variable, but the measurement is rarely made explicit. Turner and Lloyd (1999), for instance, describe exactly how they measured all their concepts, including depressive symptoms, mastery, self-esteem, chronic stressors, and life events; but they do not describe exactly how they measured depressive disorder. We show one commonly used measure in Table 1.

Using diagnosis ignores almost all information about people, treating everyone within the category (depressed) as the same, and everyone outside the category (not depressed) as the same. In reality, people who qualify for a diagnosis of depression differ in their depression levels, as do people who do not qualify for a diagnosis. This loss of information weakens correlations (Cohen 1983).

**People do not Need to be Diagnosed to be Helped**

Often the argument for categorizing people as ill versus well is that those categorized as ill can be treated. A diagnosis may or may not be handy, but it is not necessary. Anyone who feels very depressed and seeks treatment or is referred for treatment can be treated for depression. We do not need to label people as depressives, schizophrenics, or alcoholics in order to recognize that they feel bad, or their thoughts are disorganized and bizarre, or they have problems with alcohol. Certainly, we need to assess the type and extent of a person’s problems, but the assessment does not need to be categorical. A person does not have to be diagnosed to be helped. Just the opposite may
be true. Once a person receives a label, such as "schizophrenic," the diagnosis is treated as if it were the person's preeminent trait. Often the rest of the person's life and their other psychological problems are ignored. Clinical social workers use official psychiatric diagnoses primarily for insurance purposes, but they find the diagnoses of little or no value for understanding clients' psychological problems or the origins of these problems in family or work life, for understanding or predicting clients' behavior, or for planning treatment (Kirk and Kutchins 1992; Kutchins and Kirk 1988).

The Proliferation of Diagnostic Categories

Even when types of symptoms are empirically distinct, it does not mean that we can neatly assign individuals to a set of mutually exclusive diagnostic categories, saying some are depressed, others are anxious, and others are schizophrenic. Attempts to produce a set of exhaustive and mutually exclusive diagnostic categories lead to a proliferation of diagnoses that describe people who happen to have symptoms from more than one cluster. Thus, we get diagnostic categories like "schizoaffective," which is given when the clinician can't decide whether to diagnose schizophrenia (disorganized and bizarre thoughts and perceptions) or affective disorder (severe depression and anxiety). Worse than the introduction of unnecessary complexity, such a practice may obscure the fact that the causes of some symptoms on which a diagnosis is based are different than the causes of other symptoms on which the diagnosis is based.

Even with a profusion of diagnoses for people who are between categories (schizo-affective disorder), or just outside a category (schizophreniform disorder), or have an atypical disorder (atypical depression), or a mild disorder (dysthymia), a large minority of patients cannot be classified (Srole and Fischer 1980), and many problems occur together. In the National Comorbidity Survey, atypical major depression was also associated with conduct disorder, social phobia, interpersonal dependence, and alcohol/drug use disorder (Sullivan, Kessler, and Kendler 1998). Even with the profusion of diagnoses, the odds of qualifying for a diagnosis of schizophrenia are 28.5 times greater for those who also qualify for a diagnosis of major depression than for those who do not (Boyd et al. 1984).

Psychological problems are not discrete. Diagnostic categories do not reflect the reality of psychological problems. Indexes are superior scientifically to diagnostic categories, but continua are not as useful in convincing the public, other physicians, insurance companies, or government agencies that psychiatric problems are real, serious problems that deserve insurance coverage and funding (Wilson 1993). American psychiatry had an interest in developing diagnostic measures (Horwitz 2002).

A Short History of Diagnostic Instruments

When researchers in psychiatric epidemiology embarked on the development of diagnostic measures of psychological problems for use in community studies, they were faced with a problem. Diagnoses were not standardized. Clinicians making diagnoses of major depression, for example, used different criteria and made different decisions (Wakefield 1999). There was nothing obvious in what constituted the boundaries between one disorder and another, or between psychiatric illness and wellness. In reality, depression shades into anxiety, which shades into thought problems; people have more than one type of problem; and people have symptoms that range along a continuum of severity rather than being entirely present or entirely absent (Mirowsky and Ross 1983b; Wilson 1993). It is not obvious from clinical work with people who have psychological problems how to make a standardized diagnosis in which some people are categorized as depressed and others not depressed (Kutchins and Kirk 1986). The crisp "syndromes" described in diagnostic manuals are not distinct and characteristic collections of symptoms and signs that immediately impress themselves on the minds of clinical observers (Aronowitz 2001). They are official classifications laboriously negotiated over decades, and promulgated by the American Psychiatric Association (Horwitz 2002). Before the push to institute official standards in the late 1970s, there was little agreement on concepts, definitions, and diagnostic criteria (Mirowsky 1990; Wilson 1993).

The first attempts to develop standardized diagnostic instruments collected information
on the symptoms of patients in psychiatric treatment. Factor analyses of the data revealed problems that clustered empirically, but they did not correspond to clinical diagnoses (Endicott and Spitzer 1972; Spitzer et al. 1967; Spitzer et al. 1970). Rather than accept the implications of their results, psychiatric epidemiologists abandoned factor analysis as a means of shaping or validating diagnostic categories. They shifted to demonstrating "procedural validity" (Robins 1986; Robins et al. 1979), or consensus among professionals, measured as inter-rater reliability rather than internal consistency (alpha reliability). When the National Institutes of Mental Health (NIMH) contracted to develop a diagnostic protocol for use in large, multi-city surveys called the Ecological Catchment Area (ECA) studies, they considered four research diagnostic protocols. Only one had been developed and validated using factor analysis. It was immediately eliminated from any further consideration because "its scales were based on internal consistency rather than approximating traditional clinical syndromes" (Robins 1986:415). When empirical results did not conform to psychiatric preconceptions, psychiatric epidemiology abandoned the results.

Ultimately, the ECA studies used a diagnostic instrument called the Diagnostic Interview Schedule (DIS), which evolved from the Schedule for Affective Disorders and Schizophrenia (SADS), which we will describe below.

We will describe the transition from a situation of little consensus about diagnoses in the sixties and seventies to the fully reified diagnostic system of today. The detailed examples will show how this system was constructed.

Although diagnostic instruments vary somewhat, all are based on descriptions in the Diagnostic and Statistical Manual of Mental Disorders (DSM III-R and IV) of the American Psychiatric Association (1987, 1994). The general criteria for diagnosis are the presence of symptoms, co-existence of symptoms, and prolonged duration of the symptoms. Each criterion has a cut-off point below which the person does not qualify for a diagnosis of a particular disorder. There are also exclusion criteria. For example, depression with a medical cause, such as infection, anemia, or life-threatening disease, is excluded from the diagnostic category of depression, as is depression caused by a death in the family. Earlier instruments also included impaired functioning as a criterion, but later ones do not. Earlier instruments considered diagnostic categories to be mutually exclusive, so that alternate diagnoses had to be ruled out, but later ones allow for multiple diagnoses, which they call comorbidity (see, for example, Kessler and Zhao 1999).

We describe an earlier instrument called the SADS/RDC and a later one called the DIS because comparison of the two shows the evolution of diagnosis. The early instruments show the decision rules more explicitly; the later ones obscure them. There are four steps used in making a diagnosis using the SADS/RDC. The first is assessing the level of symptoms, and duration of problems. Note that level, extent, and duration all refer to assessments of degree or amount. The second is splitting each assessed amount at some cutoff point, so that differences in degree are collapsed into two categories: amounts that meet the criterion and amounts that do not. The third is totaling up so that all possible combinations of met/unmet on the three criteria are represented in a single overarching split. The fourth is excluding cases that also meet other criteria considered preeminent (such as recent bereavement).

The later instrument, the DIS, combines the first two steps. It skips the assessment of symptom levels. Instead, the split is built into the question. All questions are asked in a yes/no format, such as, "Have you ever had a period of two weeks or more when you had trouble falling asleep, staying asleep or with waking up too early? Yes or No." Respondents do not report how frequently they had trouble falling asleep, how long the period lasted, or when it happened. Thus, the information is not assessed and then ignored, as in earlier instruments; it is never assessed (Mirowsky and Ross 1989a).

Diagnosis combines assessment with judgment. The earlier instruments make this clear. Questioning, observing, and recording symptoms, functioning, and duration is assessment. Using the answers, observations, and records to assign a case to a category is judgment. The two kinds of tasks can be divided in time or between actors (e.g., nurse and physician). In the SADS/RDC they are divided into a questionnaire or protocol for assessment and an "algorithm" or set of rules for making a judg-
ment. Later instruments combine the two, going right to the judgment.

**Diagnosing Major Depression**

To illustrate psychiatric diagnosis, we'll see how the two diagnostic instruments are used to diagnose major depression. The Schedule for Affective Disorders and Schizophrenia (SADS) is used to make a diagnosis based on the Research Diagnostic Criteria (RDC) (Endicott and Spitzer 1979; Spitzer and Endicott 1978). The SADS is a questionnaire and protocol for assessing symptoms, functioning, and duration. The RDC is a set of criteria for deciding on a diagnosis given the information in the SADS. Together these are often called the SADS/RDC. Later it was replaced by the Diagnostic Interview Schedule (DIS) (Robins et al. 1979; Robins 1986). Both instruments were developed for research, and both have been used in community surveys using trained interviewers who are not psychiatrists (Weissman and Myers 1978), including large community surveys done in various metropolitan areas throughout the United States, called Epidemiologic Catchment Area (ECA) surveys (Eaton et al. 1986). Somewhat later, the National Comorbidity Survey used a modification of the DIS called the Composite International Diagnostic Interview (CIDI) (Robins et al. 1988; Kessler and Zhao 1999). The logic of diagnostic instruments remained the same; the major difference is that the CIDI is an international instrument, so that it is now used throughout the world. (For brief descriptions of other similar diagnostic instruments see Switzer, Dew, and Bromet 1999).

The earlier instrument, the SADS/RDC, is more standardized than the Diagnostic and Statistic Manual (DSM), and it is more explicit and detailed in describing how a diagnosis is made than the later research diagnostic instruments, the Diagnostic Interview Schedules (DIS and its successor, CIDI), but the SADS/RDC, the DSM, and the DIS have common roots (Horwitz 2002). Spitzer and Endicott based the RDC on the Feighner criteria (Feighner et al. 1972). As head of the task force to develop the DSM-III, Spitzer based the DSM-III on the RDC to the extent that committee decision-making would allow (Robins 1986). The DIS, developed for the ECA studies, also has its roots in the Feighner criteria and the diagnostic instrument used to make diagnoses based on these criteria, called the Renard Diagnostic Interview.

The SADS interviewer first assesses dysphoric (or depressed) mood on a continuum by asking the respondent, “How have you been feeling? Describe your mood. Have you felt depressed (sad, blue, moody, down, empty, hopeless, as if you didn’t care)? How often? Does it come and go? How long does it last? How bad is the feeling?” Based on the person’s response, the interviewer records the frequency and severity of depressed mood from not at all depressed to constant, unrelenting, extremely painful feelings of depression. Similar scales record the frequency and intensity of self-reproach, feelings of inadequacy, suicidal tendencies, sleep problems, trouble concentrating, loss of interest or pleasure, loss of energy, and other elements of depression (Spitzer and Endicott 1978).

After the questions about specific symptoms, the interviewer makes a global assessment of the subject’s level of functioning. The assessment ranges from “good functioning in all areas, many interests, socially effective,” to “major impairment in several areas such as work or family relations,” on up to “needs constant supervision for several days to prevent hurting self or others, or makes no attempt to maintain minimal personal hygiene.”

Once the assessments are made and recorded, the Research Diagnostic Criteria are applied. The symptom criterion has two conditions, which must both be met: (1) dysphoric mood characterized by feeling depressed, blue, sad, hopeless, irritable, and (2) at least five of the following eight symptoms: poor appetite or increased appetite, sleep problems, loss of energy, psychomotor retardation or agitation, loss of interest or pleasure in usual activities, self-reproach, diminished ability to concentrate, and thoughts of suicide. Although the SADS assessment records the frequency and intensity of each symptom, the criteria only refer to the presence or absence of symptoms. The duration criterion is that dysphoric mood has lasted at least one week. The functioning criterion is that the person sought or was referred for help, took medication for the problem, or had impaired functioning with family, at home, at school, at work, or socially. Again, although the SADS assessment records the frequency and intensity of dysfunction, the criterion only refers to its presence. If all three cri-
teria are met, if the person is not bereaved, if the person is not suffering from a medical illness that could cause the symptoms, and if there are no signs of schizophrenia, the person is diagnosed as having major depression.

The example illustrates the fact that diagnosing current major depression was a two-part process of collecting information and then ignoring most of it. Information is thrown out in the splitting, toting, and excluding processes. Splitting dispenses with much of it. Differences in the frequency and intensity of each symptom are ignored by counting only the presence or absence of each. Information on the number of symptoms is thrown out by ignoring less than 5, and treating any number greater than or equal to 5 as alike. Information on the duration of symptoms is thrown out by ignoring episodes of less than a week, and treating all episodes of more than a week as alike. Differences in the type of dysfunction are ignored by lumping them together, and differences in the extent of dysfunction are ignored by counting only the presence or absence of some sign of dysfunction.

The toting and excluding processes throw out information by ignoring the distinction between an emotional state and its causes and consequences (Mirowsky and Ross 1989a). When symptoms, functioning, and duration are toed up into one global judgement, the distinctions among the three dimensions are ignored, as well as the distinctions within each. For example, the practice of excluding depression caused by loss of a loved one from the diagnosis of depression obscures social patterns. Exclusion rules ignore the distinction between the emotional state and its cause or concomitant. The rules typically presuppose that each distinct cause necessarily produces a distinct disorder, even if the disorders are not otherwise distinguishable. Psychiatrists cannot tell the difference between “grief” and “major depression” without knowing if a patient has lost a close friend or relative. As a practical matter, it is worth noting whether a depressed person has recently lost a loved one, just as it is worth noting other possible causes of the depression. There is no reason to assume that each cause produces a unique emotional state. One consequence of excluding depression caused by medical illness or loss of a loved one is that social patterns are obscured. The elderly and poor have more illness and higher death rates among their loved ones than do the young and well to do. The emotional distress caused by this illness and loss is discounted in studies of diagnosed major depression. As an example, Boyd et al. (1982) describe the case of an 85-year-old woman who lives alone, without friends or family nearby. She is afraid of being robbed and will not leave her apartment. Her husband recently died. She is incapacitated by arthritis. She scores very high on any measure of depressive symptoms. She does not get a diagnosis of depression. Diagnostic exclusion criteria that discount depression due to death of a loved one, illness, injury, drugs or medications sometimes account for the fact that research diagnostic instruments count fewer people as depressed than do clinicians (Eaton et al. 2000), disproportionately discounting the depression of the elderly and poor (Bruce 1999).

Diagnoses of depression arbitrarily eliminate some causes of depression. No cause should be eliminated a priori. The research questions of interest are mostly about the causes of depression. Eliminating a few causes, such as death of a loved one and illness, confounds cause and effect, and arbitrarily eliminates some causes, but not others. Why not eliminate depression caused by one’s child’s sickness, or poverty, or divorce? If, through measurement exclusion rules, psychiatrists could eliminate all social causes of depression, they would have, by default, a measure whose causes were unseen and unmeasured biological or genetic causes. This appears to be a goal of some (Wakefield 1999). For the purpose is scientific investigation, however, it is better not to eliminate any purported cause, but rather to measure the hypothesized causes separately and correlate them with depression. No causes should be assumed or ruled out; all should be empirically examined.

Diagnoses also arbitrarily require the coexistence of symptoms. For example, in order to qualify for a diagnosis of depression mood must co-exist with malaise, which is the physiological component of depression. Malaise includes trouble sleeping, appetite problems, trouble concentrating, and feeling tired or run down. The relationship of mood and malaise should be an empirical question. Possibly, different ethnic groups, age groups, or genders express depression differently, with more or less of a mood or a malaise component. Correlations between mood and malaise should be examined, not assumed.
Historically, research diagnostic instruments ignored information on the level of a person's problems, instead splitting the information into a crude yes or no distinction, and they confounded information on symptoms, their causes, and their consequences; but the later diagnostic instruments make these problems worse than the earlier ones. The later diagnostic instrument used in the Epidemiologic Catchment Area (ECA) surveys and Comorbidity surveys is the Diagnostic Interview Schedule (DIS). The symptoms of depression are the same in the DIS as in the SADS (feelings of depression, loss of interest or pleasure [in things usually enjoyed, including sex], appetite problems, feelings of worthlessness, sleep problems, trouble thinking or concentrating, loss of energy or restlessness, and suicidal tendencies); but the format for assessing these problems in the DIS violates almost all psychometric principles of good measurement. The DIS questions used to diagnose major depression are shown in Table 1.

The first major difference is that the DIS combines the assessment of problems with the cutoff rules, whereas the SADS/RDC keeps them separate. The SADS first counts symptoms and then splits them into enough for a diagnosis versus not enough. The DIS goes right to the split. People are first asked, "In your lifetime, have you ever had two weeks or more during which you felt sad, blue, depressed or when you lost all interest and pleasure in things that you usually care about or enjoyed? Yes or no." The intensity and duration of the feelings are compressed into a simple yes or no. All other symptoms of depression are also recorded in a yes/no format (see Table 1). The DIS forces respondents to answer "yes" or "no" to questions about feelings that are not easily dichotomized. Likert scales more accurately assess people's feelings of mood and malaise.

The second major difference is the focus on lifetime problems, with current problems simply part of the lifetime assessment. (The SADS assesses current and lifetime problems with two distinct schedules). It is well established in survey research that people can most accurately report current or recent states. People have much more difficulty remembering things that happened years ago, and memory is heavily influenced by current feelings. Yet it is not until close to the end of the section on major depression (question 26; see Table 1) that the person is finally asked, "Are you in one of these spells of feeling low or disinterested and having some of these other problems now? (yes or no)."

Recall is a serious problem in measures that are about emotions, thoughts, or feelings, yet the DIS begins by asking to report whether they have ever been depressed in their lifetime. People tend to report the ways that they felt in the last couple of weeks, or maybe the last month. When asked if there was ever in your lifetime a time that they felt sad, depressed, and so on, people remember the most recent time they felt that way, and that is what they report. People can report more objective things about their whole life, like whether their parents got divorced, when they got married, how old they were when they had their first child, what their first job was; but reports on subjective things like feeling sad or depressed cannot be reliably reported for one's lifetime.

A third major difference between the DIS and the SADS is that the DIS does not emphasize impaired functioning as a criterion for receiving a diagnosis. In the SADS, a prominent indicator of impaired functioning is the act of seeking or being referred for treatment. Psychiatric epidemiologists recognized the problem of using this as a criterion for diagnosis: If, by definition, a person is not depressed if he or she does not seek treatment then there is no need for more psychiatric services.

A minor difference in duration criteria illustrates that any cutoff is arbitrary. Diagnostic instruments make arbitrary duration and severity cutoffs: the duration criteria for symptoms is one week in the SADS, while it is two weeks in the DIS. Why would someone who felt depressed for 10 days not be considered depressed? Why would someone who felt depressed for two weeks be considered equally depressed as someone who felt depressed for two years?

Why build in complex decision rules; dichotomous responses; frequency cutoffs, and so on? These are the same questions used in scales of depression, just concatenated in obscure ways. The measurement of the diagnosis of major depression is sometimes described as a "detailed computer algorithm" (Turner and Lloyd 1999:393), which we show in Table 1. Why not just ask people about their symptoms, and let them answer on a scale of frequency or severity? We think the answer is that diagnostic measures reify psychological prob-
lems, making the disorder of major depression seem like it is something different than feeling depressed, having trouble sleeping, trouble concentrating, not enjoying life, or feeling tired, sad, demoralized, worthless, hopeless about the future, and so on. Diagnostic instruments obscure more than they reveal.

**Diagnosing Schizophrenia**

Many people might agree that depression is not a categorical problem; that this is a reified notion of disease that does not reflect reality. But the same people might disagree about schizophrenia, arguing that the diagnostic approach is appropriate here. We do not think so. Symptoms of schizophrenia can also be measured by a continuous scale. Wheaton (1985) uses a scale of schizophrenic (cognitive) symptoms, shown in Table 2. Each symptom is scored according to how often the respondent reported it (never, rarely, sometimes, often, very often) and then added together. The symptoms are: felt that your mind was dominated by forces beyond your control; felt sure everyone was against you; heard voices without knowing where they heard them; obscure more than they reveal.

**TABLE 2. Short Indexes**

Respondents are asked, “How often in the past month have you____?” Responses are recorded as never (0), almost never (1), sometimes (2), fairly often (3), or often (4). Alternatively, respondents are asked, “Now I’m going to read a list of different feelings that people sometimes have. After each one I would like you to tell me on how many days you have felt this way during the last week. On how many days have you?” (0-7).

1. DISTRESS
   a. DEPRESSION
      i. Mood
         - Felt sad
         - Felt lonely
         - Felt you couldn’t shake the blues
         - Felt depressed
         - Been bothered by things that don’t usually bother you
         - Wondered if anything was worthwhile anymore
         - Felt that nothing turned out for you the way you wanted it
         - Felt completely hopeless about everything
         - Felt worthless
         - Thought about taking your own life
      ii. Malaise
         - Felt that everything was an effort
         - Felt you just couldn’t get going
         - Had trouble keeping your mind on what you were doing
         - Had trouble getting to sleep or staying asleep
         - Didn’t talk to anyone or talked less than usual
         - Felt no interest in anything or anybody
         - Felt tired all the time
         - Had poor appetite
      iii. Positive affect
         - Enjoyed life
         - Felt hopeful about the future
         - Felt happy
   b. ANXIETY
      i. Mood
         - Worried a lot about little things
         - Felt anxious, tense, or nervous
         - Felt restless or fidgety
      ii. Malaise
         - Had dizziness
         - Had shortness of breath when you were not exercising or working hard
         - Had your heart beating hard when you were not exercising or working hard
         - Suddenly felt hot all over
   c. ANGER
      i. Mood
         - Felt annoyed with things or people
         - Felt angry
      ii. Behavioral
         - Yelled at someone

Wheaton (1985) uses a scale of schizophrenic (cognitive) symptoms, shown in Table 2. Each symptom is scored according to how often the respondent reported it (never, rarely, sometimes, often, very often) and then added together. The symptoms are: felt that your mind was dominated by forces beyond your control; felt sure everyone was against you; heard voices without knowing where they heard them; obscure more than they reveal.
came from; had trouble thinking; believed you were being plotted against; thought people were saying all kinds of things about you behind your back; felt you had enemies who really wished to do you harm; been sure that everyone was against you.

Note: Researchers can use the level of generality or specificity relevant to their research question. At a general level, for example, a distress index can include depression and anxiety, mood and malaise, and positive affect (scored in reverse). At more specific levels, researchers can distinguish between depression and anxiety, and further between depressed (or anxious) mood and malaise; and anger. Cognitive symptoms can also be distinguished. In each case symptoms are added together into a scale or index.

\[ a \] from the short version of the Center for Epidemiological Studies Depression Scale (Modified CES-D) (Mirowsky and Ross 1990; Ross and Mirowsky 1984)
\[ b \] from the CES-D (Radloff 1977)
\[ c \] modified version of symptoms from the Langner index (Langner 1962)
\[ d \] from the schizophrenia scale (Wheaton 1985)
\[ e \] from the mistrust and paranoia scales (Mirowsky and Ross 1983a, 1983b).
\[ f \] from the anger scale (Mirowsky and Ross 1995, 1996; Ross and VanWilligen 1996).

TABLE 2. (Continued)

<table>
<thead>
<tr>
<th>2. COGNITIVE (SCHIZOPHRENIC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. DELUSIONS AND HALLUCINATIONS</td>
</tr>
<tr>
<td>Felt that your mind was dominated by forces beyond your control(^d)</td>
</tr>
<tr>
<td>Heard voices without knowing where they came from(^d)</td>
</tr>
<tr>
<td>Had visions or seen things other people say they cannot see(^d)</td>
</tr>
<tr>
<td>Felt that you were possessed by a spirit or devil(^d)</td>
</tr>
<tr>
<td>Felt you had special powers(^d)</td>
</tr>
<tr>
<td>Felt that you did not exist at all, that you were dead, dissolved(^d)</td>
</tr>
<tr>
<td>Seemed to hear your thoughts spoken aloud—almost as if someone standing nearby could hear them(^d)</td>
</tr>
<tr>
<td>Felt that your unspoken thoughts were being broadcast or transmitted, so that everyone knew what you were thinking(^d)</td>
</tr>
</tbody>
</table>

| b. PARANOIA |
| Felt it was safer to trust no one\(^e\) |
| Been very suspicious\(^e\) |
| Believed you were being plotted against\(^e\) |
| Felt that people were saying all kinds of things about you behind your back\(^e\) |
| Felt you had enemies who really wished to do you harm\(^e\) |
| Been sure that everyone was against you\(^e\) |

Implications of Measurement for Stress Research

Indexes provide advantages over diagnoses in research on social causes of psychological distress. One important advantage is that indexes do not discount symptoms that coincide with stressors that might have caused the symptoms. The more that diagnoses try to distinguish clinical cases from those caused by frustration, inequality, disadvantage, strain, and deprivation, the more they eliminate the

JOURNAL OF HEALTH AND SOCIAL BEHAVIOR
very correlations with stressful social conditions with which sociologists are concerned. Also, the more that diagnoses eliminate the predictable cases, the less information the data based on them can yield. Indexes assess the level of symptoms without discounting those that coincide with sickness, disability, loss, or other possible causes. Thus, indexes allow researchers to measure, test, and explain the correlations of those possible causes with symptoms (Mirowsky 1994).

Another important advantage of indexes for stress research is that they assess the full range of symptoms, and not just the extremes that justify clinical action. At heart, stress research provides information meant to help people avoid those extremes. Most differences in mental, emotional, and behavioral problems occur within the large majority of people not qualified for a psychiatric diagnosis (Horwitz and Scheid 1999). Those differences provide most of the information available and relevant to prevention. This is true even of problems not traditionally studied across their full spectrum, such as paranoia and schizophrenia. For example, only about 1 percent of the general population has symptoms that qualify for a clinical diagnosis of schizophrenia (Holzer et al. 1986). Thus, typical community samples of 500 to 2,000 might find about 5 to 20 persons qualified for the diagnosis, which is not enough to study. Nevertheless, even a sample of less than 500 finds scores of people who, at some time in the previous year, felt their minds dominated by outside forces, heard voices, had visions, felt possessed, felt they had special powers, felt nonexistent, heard their own thoughts as if spoken aloud, or felt their thoughts were being broadcast. Even in such a small sample, an index of schizophrenia provides rich information on the effects of stressors and of moderators such as mastery and flexibility on the commonness and intensity of these problems (Wheaton 1985).

Another advantage of indexes for stress research lies in their efficiency, precision, and specificity. One claim made by proponents of diagnostic assessment is that diagnoses distinguish different types of problems better than indexes do (e.g., Eaton and Kessler 1985). In fact, just the opposite is true. Diagnoses typically combine several distinct attributes into a single dichotomous variable. Thus, correlations with the diagnosis cannot show which diagnostic attribute is responsible for the correlation. In addition, two different diagnostic categories often share one or more defining attributes. Correlations between diagnoses reflect the proportion of attributes shared. The more that two diagnostic categories, such as schizophrenia and depression, share the same defining attributes, the less one can distinguish findings and statements about one from those about the other.

By classifying symptoms rather than people, indexes distinguish, rather than blend, different types of problems. By design, each index represents one distinct attribute. Two distinct indexes do not share elements. Thus, indexes minimize confusion about the type of mental, emotional, or behavioral problem that arises. Table 2 lists symptoms that form indexes of depression, anxiety, anger, delusions and hallucinations, and paranoia. By using these or similar indexes, stress researchers maximize statistical power and analytic clarity. The strength of an index flows from its simplicity and unity.

CONCLUSION

For researchers who decide to create knowledge that gives power to those who suffer or risk suffering, certain things follow. In one way or another they all come down to one thing: respect for the autonomy and self-determination of the people we study. Logically, then, the science we create should be designed to educate rather than to manipulate. It should use common language as much as possible. It should take measure of life as people feel it, sense it, and experience it. We connect our science to the lives of those we study by measuring the shades and grades of misery, anguish, distress, and alarm common to human beings—the human burden of tension, worry, mistrust, apprehension, fear, dejection, despair, sadness, loneliness, shame, frustration, humiliation, resentment, indignation, hostility, scorn, and anger. Although most measures have focused on the negative end of human experience, we hope to likewise connect by measuring the qualities of existence humans enjoy and seek: happiness, delight, serenity, fervor, gratification, hope, joy, pur-
pose, accomplishment, affection, and acceptance. Let the sociology of mental health observe, think, and speak in such terms. Let us create a human science.

REFERENCES


John Mirowski is Professor of Sociology at Ohio State University. In September 2002, he will join the Department of Sociology and Population Research Center at the University of Texas as a Professor. He is principal investigator on a project funded by the National Institute on Aging to study “Aging, Status and the Sense of Control” and a project funded by the National Institute of Mental Health to study “Children, Child-Care and Psychological Well-Being” (Catherine Ross is co-principal investigator on both). His recent and forthcoming publications include “Age, Depression and Attrition in the National Survey of Families and Households” Sociological Methods and Research 2000 (with John R. Reynolds), and “Parenthood and Health: The Pivotal and Optimal Age at First Birth” forthcoming Social Forces. With Catherine Ross he is working on the second edition of Social Causes of Psychological Distress, Aldine de Gruyter.

Catherine E. Ross is a Professor in the Department of Sociology at the Ohio State University. In September 2002, she will join the Department of Sociology and Population Research Center at the University of Texas as a Professor. She studies the effects of socioeconomic status, work, family and community on men's and women's physical and mental health, and their sense of control versus powerlessness. Recent publications include “Does Medical Insurance Contribute to Socioeconomic Differentials in Health.” Milbank Quarterly 2000 (with John Mirowski); “Neighborhood Disadvantage, Disorder and Health” Journal of Health and Social Behavior 2001 (with John Mirowski); and “Powerlessness and the Amplification of Threat: Neighborhood Disadvantage, Disorder, and Mistrust” American Sociological Review 2001 (with John Mirowski and Shana Pribesh). With John Mirowski she is writing a book on Education and Health.