More on Pseudoscience in Science and the Case for Psychiatric Diagnosis

A Critique of D. L. Rosenhan's "On Being Sane in Insane Places" and "The Contextual Nature of Psychiatric Diagnosis"

Robert L. Spitzer, MD

• **Rosenhan's 1973 article,[1] "On Being Sane in Insane Places," was pseudoscience presented as science. Just as his pseudopatients were diagnosed at discharge as having "schizophrenia in remission," so a careful examination of this study's methods, results, and conclusions leads to a diagnosis of "logic in remission." Rosenhan's study proves that pseudopatients are not detected by psychiatrists as having simulated signs of mental illness and that the implementation of certain invalid research designs can make psychiatrists appear foolish. These rather unremarkable findings are irrelevant to the real problems of the reliability and validity of psychiatric diagnosis and only serve to obscure them. A correct interpretation of his own data contradicts his conclusions. There are purposes to psychiatric diagnosis that Rosenhan's article ignores. His more recent suggestion that certain requirements be met prior to the adoption of a new psychiatric classification system is unrealistic.**

(*Arch Gen Psychiatry* 33:459-470, 1976)

In January 1973, *Science*, the official journal of the Association for the Advancement of Science, reported a small study with a catchy title—"On Being Sane in Insane Places.[1]" This was no ordinary study that merely added further knowledge to our understanding of psychiatric disorders; this study challenged basic psychiatric concepts and practices. If the author, D. L. Rosenhan, a professor of psychology and law, is correct, the results clearly show that psychiatrists are unable to distinguish the "sane" from the "insane" in psychiatric hospitals, and that the traditional psychiatric classification of mental disorders is unreliable, invalid, and harmful to the welfare of patients.

Partly because of the prestige of the journal in which it first appeared, and more importantly, because it said something that many were delighted to hear, the study was widely acclaimed in the popular news media (*New York Times*, Jan 20, 1974; *Saturday Review of Science*, March 1, 1973, pp 55-56; *Newsweek*, Jan 29, 1973, p 46). As a consequence, this single study is probably better known to the lay public than any other study in the area of psychiatry in the last decade.

Although the study has been attacked by many mental health professionals,[2-9] most articles that refer to the study have accepted its conclusions and implications.[10-24] Furthermore, two editorials in *The Journal of the American Medical Association* were devoted to an endorsement of the study's findings.[13,14]

The study has probably had its greatest impact in the field of psychology. Of 31 recently published psychology textbooks, 15 cite Rosenhan's article. Fully 12 of these texts[25-36] present the results uncritically with only five[27,35,37-39] even acknowledging controversy over the study's conclusions. Only three[37-39] question its results. The implication is clear: large numbers of undergraduate and graduate psychology students are being taught to accept the conclusions of this study.

Although there have been references to Rosenhan's study in articles appearing in well-known psychiatric journals, none has presented a thorough critique of this study. Such a critique would be useful, not only for what it tells us about Rosenhan's remarkable study, but also for clarifying some of the fundamental issues regarding psychiatric diagnosis. This article is such an attempt, and presents an elaboration of my contribution to a symposium that appeared in *The Journal of Abnormal Psychology*,[40-45] which was devoted to exploring the strengths and weaknesses of the Rosenhan study. In addition, this article includes a critique of Rosenhan's contribution to the symposium, an article entitled, "The Contextual Nature of Psychiatric Diagnosis,[43]" in which he responded to the critiques provided by the participants in the symposium.

Rosenhan stated the basic issue in his original article as follows:

> Do the salient characteristics that lead to diagnoses reside in the patients themselves or in the environments and contexts in which observers find them? From Bleuler, through Kretchmer, through the formulators of the recently revised *Diagnostic and Statistical Manual* of the American Psychiatric Association, the belief has been strong that patients present symptoms, that those symptoms can be categorized, and, implicitly, that the sane are distinguishable from the insane. More recently, however, this belief has been questioned. Based in part on theoretical and anthropological considerations, but also on philosophical, legal and therapeutic

Accepted for publication Dec 12, 1975.

From Biometrics Research, New York State Psychiatric Institute, and the Department of Clinical Psychiatry, Columbia University, New York.

Reprint requests to New York State Psychiatric Institute, 722 W 168th St, New York, NY 10032 (Dr Spitzer).

ones, the view has grown that psychological categorization of mental illness is useless at best and downright harmful, misleading, and pejorative at worst. Psychiatric diagnoses, in this view, are in the minds of the observers and are not valid summaries of characteristics displayed by the observed.[1(p251)]

Rosenhan proposed that an adequate method to study this question was for normal people who had never had symptoms of serious psychiatric disorders to be admitted to psychiatric hospitals "and then determining whether they were discovered to be sane." Therefore, eight "sane" people, or "pseudopatients," gained admission to 12 different hospitals, each with a single complaint of hearing voices. On admission to the psychiatric ward, each pseudopatient ceased simulating any symptoms of abnormality.

The diagnostic results were as follows:

Admitted, except in one case, with a diagnosis of schizophrenia, each was discharged with a diagnosis of schizophrenia "in remission." The label "in remission" should in no way be dismissed as a formality, for at no time during any hospitalization had any question been raised about any pseudopatient's simulation.[1(p252)]

(It should be noted that while preparing my original critique, personal communication with Rosenhan indicated that "in remission" referred to use of that term or one of its equivalents, such as "recovered" or "no longer ill," and that it also applied to the one patient who was given the diagnosis of manic-depressive psychosis. Thus, *all* of the patients were apparently discharged "in remission." However, in his 1975 article,[43] he notes that only eight of the patients were discharged "in remission" and that one was noted as "asymptomatic" and three as "improved." The discrepancy between the 1973 and 1975 articles is puzzling but does not substantially alter my interpretation of the results.)

Rosenhan concluded, "It is clear that we cannot distinguish the sane from the insane in psychiatric hospitals."[1(p257)] According to him, what is needed is the avoidance of "global diagnosis" as exemplified by such diagnoses as schizophrenia or manic-depressive psychosis, and attention should be directed instead to "behaviors, the stimuli that provoke them, and their correlates."

THE CENTRAL QUESTION

One hardly knows where to begin. Let us first acknowledge the potential importance of the study's central research question. Surely, if psychiatric diagnoses are, to quote Rosenhan, "only in the minds of the observers," and do not reflect any characteristics inherent in the patient, then they obviously can be of no use in helping patients. (It is remarkable that the original article, which was concerned with the validity of psychiatric diagnosis, did not contain a single sentence about the intended purposes of psychiatric diagnosis—more of this later.) However, the study immediately becomes hopelessly confused when Rosenhan suggests that his research question can be answered by studying whether or not the "sanity" of pseudopatients in a mental hospital can be discovered. Rosenhan, a professor of law and psychology, knows that the terms "sane" and "insane" are legal, not psychiatric, concepts. He knows that no psychiatrist makes a *diagnosis* of "sanity" or "insanity," and that the true meaning of these terms, which varies from state to state, involves the inability to distinguish right from wrong—an issue that is totally irrelevant to this study.

DETECTING THE SANITY OF A PSEUDOPATIENT

However, if we are forced to use the terms "insane" (to mean showing signs of serious mental disturbance) and "sane" (the absence of such signs), then clearly there are three possible meanings to the concept of "detecting the sanity" of a pseudopatient who feigns mental illness on entry to a hospital but then acts "normal" throughout his hospital stay. The first is the recognition, *when he is first seen*, that the pseudopatient is feigning insanity as he attempts to gain admission to the hospital. This would be detecting sanity in a sane person simulating insanity. The second would be the recognition, *after* having observed him acting normally during his hospitalization, that the pseudopatient was initially feigning insanity. This would be detecting that the currently sane person never was insane. Finally, the third possible meaning would be the recognition, *during hospitalization*, that the pseudopatient, though initially appearing to be "insane," was no longer showing signs of psychiatric disturbance.

These elementary distinctions of "detecting sanity in the insane" are crucial to properly interpreting the results of Rosenhan's study. The reader is misled by Rosenhan's implication that the first two meanings of detecting the sanity of the pseudopatients, which involve determining the pseudopatient to be a fraud, are at all relevant to the central research question. Further, the true results of his study are obscured because they fail to support the conclusion when the third meaning of detecting sanity is considered, that is, a recognition that after their admission as "insane," the pseudopatients were not psychiatrically disturbed while in the hospital.

Let us examine these three possible meanings of detecting the sanity of the pseudopatient, their logical relation to the central question of the study, the actual results obtained, and the validity of Rosenhan's conclusions.

THE PATIENT IS NO LONGER "INSANE"

We begin with the third meaning of detecting sanity. It is obvious that if the psychiatrists judged the pseudopatients as seriously disturbed while they acted "normal" in the hospital, this would be strong evidence that their assessments were being influenced by the context in which they were making their examination rather than the actual behavior of the patient. This, after all, is the central research question. (I suspect that many readers will agree with Hunter, who, in a letter to *Science*, pointed out:

The pseudopatients did *not* behave normally in the hospital. Had their behavior been normal, they would have walked to the nurses' station and said, "Look, I am a normal person who tried to see if I could get into the hospital by behaving in a crazy way or saying crazy things. It worked and I was admitted to the hospital, but now I would like to be discharged from the hospital.[46])

What were the results? According to Rosenhan, all of the patients were diagnosed at discharge as being "in remission." The meaning of "in remission" is obvious: it means without signs of illness. Thus, the psychiatrists apparently recognized that the pseudopatients were, to use Rosen-

han's term, "sane." (This would apply to all of the 12 pseudopatients according to the 1973 article and to eight of them according to the 1975 article.) However, lest the reader appreciate the significance of these findings, Rosenhan gives a completely incorrect interpretation: "If the pseudopatient was to be discharged, he must naturally be 'in remission'; but he was not sane, nor, in the institution's view, had he ever been sane."[1(p252)] Rosenhan's implication is clear—the patient was diagnosed "in remission" not because the psychiatrist correctly assessed the patient's hospital behavior, but only because the patient had to be discharged. Is this interpretation warranted?

I am sure that most readers who are not familiar with the details of psychiatric diagnostic practice assume from Rosenhan's account that it is common for schizophrenic patients to be diagnosed "in remission" when discharged from a hospital; as a matter of fact, this is extremely unusual. The reason is two-fold. First of all, patients with a diagnosis of schizophrenia are rarely completely asymptomatic at discharge. Second, the discharge diagnosis frequently records the diagnostic conditions associated with the *admission* to the hospital without any reference to the condition of the patient at discharge.

Rosenhan does not report any data concerning the discharge diagnoses of the real schizophrenic patients in the 12 hospitals used in his study. However, I can report on the frequency of a discharge diagnosis of schizophrenia "in remission" at my hospital, the New York State Psychiatric Institute, a research, teaching, and community hospital where diagnoses are made in a routine fashion, undoubtedly no differently from the 12 hospitals of Rosenhan's study. I examined the official book that the record room uses to record discharge diagnoses and their statistical codes for all patients. Of more than 300 patients discharged in the year prior to September 1974 with a diagnosis of schizophrenia, not one was diagnosed "in remission." It is only possible to *code* a diagnosis of "in remission" by adding a fifth digit (5) to the four-digit code number for the subtype of schizophrenia (eg, paranoid schizophrenia is coded as 295.3, but paranoid schizophrenia "in remission" is coded as 295.35). I realize, however, that a psychiatrist might *intend* to make a discharge diagnosis of "in remission" but fail to use the fifth digit, so that the official recording of the diagnosis would not reflect his full assessment. I therefore had research assistants read the discharge summaries of the last 100 patients whose discharge was schizophrenia to see how often the terms "in remission," "recovered," "no longer ill," or "asymptomatic" were used, even if not recorded with the fifth digit in the code number. The result was that only one patient, who was diagnosed paranoid schizophrenia, was described in the summary as being "in remission" at discharge. The fifth digit code was not used.

To substantiate my view that the practice at my hospital of rarely giving a discharge diagnosis of schizophrenia "in remission" is not unique, I had a research assistant call the record room librarians of 12 psychiatric hospitals, chosen "catch-as-catch-can." (Rosenhan explains his refusal to identify the 12 hospitals used in his study on the basis of his concern with issues of confidentiality and the potential for ad hominem attack. However, this makes it impossible for anyone at those hospitals or elsewhere to corroborate or challenge his account of how the pseudopatients acted and how they were perceived.) The 12 hospitals used in my ministudy were the following: Long Island Jewish-Hillside Medical Center, New York; Massachusetts General Hospital, Massachusetts; St. Elizabeth's Hospital, Washington, DC; McLean Hospital, Massachusetts; UCLA, Neuropsychiatric Institute, California; Meyer-Manhattan Hospital (Manhattan State), New York; Vermont State Hospital; Medical College of Virginia; Emory University Hospital, Georgia; High Point Hospital, New York; Hudson River State Hospital, New York; and New York Hospital-Cornell Medical Center, Westchester Division. The record room librarians were told that we were interested in knowing their estimate of how often, at their hospitals, schizophrenics were discharged "in remission" (or "no longer ill" or "asymptomatic"). The results were that 11 of the 12 hospitals indicated that the term was either never used or, at most, was used for only a handful of patients in a year. The remaining hospital (a private one) estimated that the term was used in roughly 7% of the discharge diagnoses.

This leaves us with the conclusion that the pseudopatients were given a discharge diagnosis (All 12 of them? Eight of them?) that is rarely given to real patients with an admission diagnosis of schizophrenia. Therefore, the diagnoses given to the pseudopatients *were* a function of the patients' behaviors and not of the setting (psychiatric hospital) in which the diagnoses were made. In fact, a moment's reflection may cause many a reader familiar with usual diagnostic practice to marvel that so many psychiatrists acted so rationally as to use at discharge precisely the same diagnostic category, "in remission," that is rarely used with real patients. In any case, the data as reported by Rosenhan contradict his conclusions.

It is not only in his discharge diagnosis that the psychiatrist had an opportunity to assess the patient's true condition incorrectly. In the admission mental status examination, during a progress note or in his discharge note, that psychiatrist could have described any of the pseudopatients as "still psychotic," "probably still hallucinating but denies it now," "loose associations," or "inappropriate affect." Because Rosenhan had access to all of this material, his failure to report such judgments of continuing serious psychopathology, either in the original study or in his 1975 symposium article,[43] strongly suggests that they were never made.

All pseudopatients took extensive notes publicly to record data on staff and patient behavior. Rosenhan claimed that the nursing records indicated that "the writing was seen as an aspect of their pathological behavior."[1(p253)] The only datum presented to support this claim is that the daily nursing comment on one of the pseudopatients was "patient engages in writing behavior." Because nursing notes frequently and intentionally comment on nonpathological activities that a patient engages in so that other staff members have some knowledge of how the patient spends his time, this particular nursing note in no way supports Rosenhan's thesis. Once again, the *failure* to provide data regarding instances where normal hospital behavior was categorized as pathological is remarkable. The closest that Rosenhan comes to providing such data is

his report of an instance where a kindly nurse asked if a pseudopatient, who was pacing the long hospital corridors because of boredom, was "nervous." It was, after all, a question and not a final judgment.

Let us now examine the other two meanings of detecting sanity in the pseudopatients, that is, the recognition that the pseudopatient was a fraud either when he sought admission to the hospital or during his hospital stay, and the relationship of those meanings to the central research question.

DETECTING "SANITY" BEFORE ADMISSION

Whether or not psychiatrists are able to detect individuals who feign psychiatric symptoms is an interesting question, but it is clearly of no relevance to the issue of whether or not the salient characteristics that lead to diagnoses reside in the patient's behavior or in the minds of the observers. After all, a psychiatrist who believes a pseudopatient who feigns a symptom *is* responding to the pseudopatient's *behavior*. Rosenhan does not blame the psychiatrist for believing the pseudopatient's fake symptom of hallucinations. He blames him for making the diagnosis of schizophrenia. He states:

> The issue is not . . . that the psychiatrist believed him. . . . Neither is it whether the pseudopatient should have been admitted to the psychiatric hospital in the first place. . . . The issue is the diagnostic leap that was made between the single presenting symptom, hallucinations, and the diagnosis, schizophrenia (or, in one case, manic-depressive psychosis). . . . Had the pseudopatients been diagnosed "hallucinating" there would have been no further need to examine the diagnostic issue. The diagnosis of hallucinations implies only that: no more. The presence of hallucinations does not itself define the presence of "schizophrenia." And schizophrenia may or may not include hallucinations.[47(pp 366,367)]

Let us see. Unfortunately, as judged by many of the letters to *Science* commenting on the study,[6] many readers, including psychiatrists, accepted Rosenhan's thesis that it was irrational for the psychiatrists to have made an initial diagnosis of schizophrenia *as the most likely condition* on the basis of a single symptom. In my judgment, these readers were wrong. Their acceptance of Rosenhan's thesis was aided by the content of the pseudopatients' auditory hallucinations, which were voices that said "empty," "hollow," and "thud." According to Rosenhan, these symptoms were chosen because of "their apparent similarity to existential symptoms [and] the *absence* of a single report of existential psychoses in the literature."[1(p251)] The implication is that if the *content* of specific symptoms has never been reported in the literature, then a psychiatrist should somehow know that the *symptom* has no diagnostic significance. This is absurd. Recently I saw a patient who kept hearing a voice that said. "It's OK. It's OK." I know of no such report in the literature. So what? I agree with Rosenhan that there has never been a report of an "existential psychosis." However, the diagnoses made were schizophrenia and manic-depressive psychosis, not existential psychosis. (I am reminded of a game that was played when I was a kid. "I can prove that you are not here." "How?" "Are you in Chicago?" "No." "Then you must be in some other place. If you are in some other place then you can't be here.")

DIFFERENTIAL DIAGNOSIS OF AUDITORY HALLUCINATIONS

Rosenhan is entitled to believe that pschiatric diagnoses are of no use and therefore should not have been given to the pseudopatients. However, it makes no sense for him to claim that *within* a diagnostic framework it was irrational to consider schizophrenia seriously as the most likely condition *without* his presenting a consideration of the differential diagnosis. Let me briefly give what I think is a reasonable differential diagnosis, based on the initial clinical picture of the pseudopatient when he applied for admission to the hospital.

Rosenhan says that "beyond alleging the symptoms and falsifying name, vocation, and employment, no further alterations of person, history, or cirumstances were made."[1(p251)] However, the clinical picture clearly includes not only the symptom (auditory hallucinations) but also the desire to enter a psychiatric hospital, from which it is reasonable to conclude that the symptom is a source of significant distress. (In fact, in his 1975 symposium article Rosenhan acknowledges that the pseudopatients reported that "the hallucinations troubled them greatly at the outset."[43(p471)] This, plus the knowledge that the auditory hallucinations were reported to be of three weeks' duration (D. L. Rosenhan, oral communication), establishes the hallucinations as significant symptoms of psychopathology as distinguished from so-called "pseudohallucination" (hallucinations while falling asleep or awakening from sleep, or intense imagination with the voices heard from inside of the head).

Auditory hallucinations can occur in several kinds of mental disorders. The absence of a history of alcohol, drug abuse, or some other toxin, the absence of any signs of physical illness (such as high fever), the absence of evidence of distractibility, impairment in concentration, memory or orientation, and negative results from a neurological examination all make an organic psychosis extremely unlikely. The absence of a recent precipitating stress rules out a transient situational disturbance of psychotic intensity or (to use a nonofficial category) hysterical psychosis. The absence of a profound disturbance in mood rules out an affective psychosis (we are not given the mental status findings for the patient who was diagnosed manic-depressive psychosis).

What about simulating mental illness? Psychiatrists know that occasionally an individual who has something to gain from being admitted to a psychiatric hospital will exaggerate or even feign psychiatric symptoms. This is a genuine diagnostic problem that psychiatrists and other physicians occasionally confront and is called "malingering." However, there was certainly no reason to believe that any of the pseudopatients had anything to gain from being admitted to a psychiatric hospital except relief from their alleged complaint, and therefore there was no reason to suspect that the illness was feigned. What possible diagnoses are left in the classification of mental disorders now used in this country for a patient with a presenting symptom of hallucinations, with the previously considered conditions having been ruled out? There is only one—schizophrenia!

Admittedly, there is a hitch to a definitive diagnosis of schizophrenia. Almost invariably there *are* other signs of the disorder present, such as poor premorbid adjustment, affective blunting, delusions, or signs of thought disorder. I would hope that if I had been one of the 12 psychiatrists presented with such a patient, I would have been struck by the lack of other signs of the disorder, but I am rather sure that having no reason to doubt the authenticity of the patient's claim of auditory hallucinations, I also would have beel fooled into noting schizophrenia as *the most likely* diagnosis.

What does Rosenhan really mean when he objects to the diagnosis of schizophrenia because it was based on a "single symptom"? Does he believe that there are real patients with the single symptoms of auditory hallucinations who are misdiagnosed as schizophrenic when they actually have some other condition? If so, what is the nature of that condition? Is Rosenhan's point that the psychiatrists should have used "diagnosis deferred," a category that is available but rarely used? I would have no argument with this conclusion. Furthermore, if he had presented data from real patients indicating how often patients are erroneously diagnosed on the basis of inadequate information and what the consequences are, it would have been a real contribution.

Until now, I have assumed that the pseudopatients presented only one symptom of psychiatric disorder. Actually, we know very little about how the pseudopatients presented themselves. What did the pseudopatients say when asked, as most must have been, what effect the hallucinations were having on their lives? Did any of the pseudopatients depart from the protocol (which called for describing only one symptom), perhaps in an effort to justify admission to the hospital? (It occurred to me that the best way to shed light on this question would be to read the original admission notes to determine just how the psychiatrist described the present illnesses of pseudopatients. Communication with Rosenhan indicated that he has this material. I have made several requests for him to send me copies, with deletion of all information that could possibly identify the particular hospitals that were involved. To summarize a lengthy correspondence, he has indicated that editing this material is more difficult than I would judge and that he would be glad to supply the material after he has completed analyzing it for a book he is preparing.)

DETECTING SANITY AFTER ADMISSION

Let us now examine the relationship to the central research question of the last meaning of detecting sanity in the pseudopatients, namely, the psychiatrist's recognition, *after* observing the pseudopatient act normally during his hospitalization, that he was initially feigning insanity. If a diagnostic condition were known to be always chronic and unremitting, it would be irrational not to question the original diagnosis if a patient were later found to be asymptomatic. As applied to this study, if the concept of schizophrenia did not admit the possibility of recovery, then failure to question the original diagnosis when the pseudopatients were no longer overtly ill would be relevant to the central research question. It would be an example of the context of the hospital environment influencing the psychiatrist's diagnostic decision. However, neither any psychiatric textbook nor the American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders*[48] suggests that mental illnesses endure forever. Oddly enough, it is Rosenhan who, without any reference to the psychiatric literature, says, "A broken leg is something one recovers from, but mental illness allegedly endures forever."[1(p254)] Who, other than Rosenhan, alleges it?

Rosenhan should know that although some American psychiatrist restrict the label of schizophrenia to mean chronic or process schizophrenia, most American psychiatrists include an acute subtype from which there is often remission. The *Diagnostic and Statistical Manual* in describing the subtype, "acute schizophrenic episode," states that "in many cases the patient recovers within weeks."[48]

A similar straw man is created when Rosenhan says:

> The insane are not always insane ... the bizarre behaviors upon which their [the pseudopatients'] diagnoses were allegedly predicated constituted only a small fraction of their total behavior. If it makes no sense to label ourselves permanently depressed on the basis of an occasional depression, then it takes better evidence than is presently available to label all patients insane or schizophrenic on the basis of bizarre behaviors or cognitions.[1(p254)]

Who ever said that the behaviors that indicate schizophrenia or any other diagnostic category comprise the totality of a patient's behavior? A diagnosis of schizophrenia does not mean that *all* of the patient's behavior is schizophrenic, any more than a diagnosis of carcinoma of the liver means that all of the patient's body is diseased. (While discussing the pitfalls of generalizing, how about Rosenhan's conclusion that "It is clear that we cannot distinguish the sane from the insane in psychiatric hospitals,"[1(p257)] which is based on a sample size of eight pseudopatients admitted to 12 hospitals!)

Does Rosenhan at least score a point by demonstrating that, although the professional staff never considered the possibility that the pseudopatient was a fraud, this possibility was often considered by other patients? Perhaps, but I am not so sure. Let us not forget that all of the pseudopatients "took extensive notes publicly." Obviously, this was highly unusual patient behavior and Rosenhan's quote from a suspicious patient suggests the importance it had in focusing the other patients' attention on the pseudopatients: "You're not crazy. You're a journalist, or a professor [referring to the continual note-taking]. You're checking up on the hospital."[1(p252)]

Rosenhan presents ample evidence, which I find no reason to dispute, that the professional staff spent little time actually with the pseudopatients. The note-taking may easily have been overlooked, and therefore the staff developed no suspicion that the pseudopatients had simulated illness to gain entry into the hospital. The note-taking, in which all of the pseudopatients engaged, may well have been *the* cue that alerted the patients to the possibility that the pseudopatients were there under false pretenses. However, I would predict that a pseudopatient on a ward of patients with mixed diagnostic conditions would have no difficulty in masquerading convincingly as a

true patient to both staff and patients if he did nothing unusual to draw attention to himself.

Rosenhan presents one way in which the diagnosis did affect the psychiatrist's preception of the patient's circumstances—historical facts of the case were often distorted by the staff to achieve consistency with psychodynamic theories. Here, for the first time, I believe Rosenhan has hit the mark. What he described happens all the time and often makes attendance at clinical case conferences extremely painful, especially for those with logical minds and research orientations. Although his observation is correct, it would seem to be more a consequence of individuals attempting to rearrange facts to comply with an unproven etiological theory than a consequence of diagnostic labeling. One could easily imagine a similar process occurring when a weak-minded, behaviorally oriented clinician attempts to rewrite the patient's history to account for "hallucinations reinforced by attention paid to patient by family members when patient complains of hearing voices." Such is the human condition.

One final finding requires comment. In order to determine whether "the tendency toward diagnosing the sane insane could be reversed," the staff of a research and teaching hospital was informed that at some time during the following three months, one or more pseudopatients would attempt to be admitted. No such attempt was actually made. Yet approximately 10% of 193 real patients were suspected by two or more staff members (we are not told how many made judgments) to be pseudopatients. Rosenhan concluded, "Any diagnostic process that lends itself so readily to massive errors of this sort cannot be a very reliable one."[1(p252)] My conclusion is that this experimental design practically assures only one outcome. (Did the hospital director, or whoever it was that agreed to participate in this ministudy, really believe that the design was relevant to some serious research question?)

ELEMENTARY PRINCIPLES OF RELIABILITY OF CLASSIFICATION

Some very important principles that are relevant to the design of Rosenhan's study are taught in elementary psychology courses and should not be forgotten. One of them is that a measurement or classification procedure is not reliable or unreliable *in itself* but only in its application to a specific population. There *are* serious problems with the reliabiltiy of psychiatric diagnoses as they are applied to the population to which psychiatric diagnoses are *ordinarily* given. However, I fail to see, and Rosenhan does not even attempt to show, how the reliability of psychiatric diagnoses applied to a population of individuals seeking help is at all relevant to the reliability of psychiatric diagnoses applied to a population of pseudopatients (or one including the threat of pseudopatients). The two populations are just not the same. Kety has expressed it dramatically:

> If I were to drink a quart of blood and, concealing what I had done, come to the emergency room of any hospital vomiting blood, the behavior of the staff would be quite predictable. If they labeled and treated me as having a bleeding peptic ulcer, I doubt that I could argue convincingly that medical science does not know how to diagnose that condition.[3(p959)]

(I have no doubt that if the condition known as "pseudopatient" ever assumed epidemic proportions among admissions to psychiatric hospitals, psychiatrists would in time become adept at identifying them, though at what risk to real patients I do not know.)

ATTITUDES TOWARD THE "INSANE"

The latter part of Rosenhan's study[1] deals with the experience of psychiatric hospitalization. The staff and the patients were strictly segregated. The professional staff, especially the psychiatrists, were not available and were rarely seen. When the staff was asked for information "their most common response consisted of either a brief response to the question, offered while they were 'on the move' and with head averted, or no response at all."[1(p255)] "Attendants delivered verbal and occasionally serious physical abuse to patients in the presence of other observing patients."[1(p256)] One attendant awakened patients with "Come on you m-----f-----s, out of bed!"[1(p256)] "One patient was beaten in the presence of other patients for having approached an attendant and told him, 'I like you.' "[1(p256)]

Because some of the hospitals participated in residency training programs and are described as "research oriented," I do find it hard to believe that the conditions were quite as bad as depicted. Perhaps they were. But how are we then to understand Rosenhan when in the summary to his original article he says:

> It could be a mistake, and a very unfortunate one, to consider that what happened to us derived from malice or stupidity on the part of the staff. Quite the contrary, our overwhelming impression of them was of people who really cared, who were committed and who were uncommonly intelligent.[1(p257)]

Surely what he described, including the verbal and physical abuse given to patients, is hardly what most people would regard as the behavior of people who "really cared" and were "uncommonly intelligent."

There is an obvious reason for the discrepancy between the actual behavior of the staff that Rosenhan describes and his exoneration of them for any responsibility for "malice" or "stupidity." To direct attention to any shortcomings on the part of the staff would detract attention from the real culprit, namely, diagnostic labels. Thus, Rosenhan asserts, without a shred of evidence from his study, that "Negative attitudes [toward psychiatric patients]... are the natural offspring of the labels patients wear and the places in which they are found."[1(p254)] Nonsense! This makes as much sense as asserting that the attitude of the public toward cancer is the natural offspring of the label "cancer" without considering the attitude of the public to any of the *features* of neoplastic disease.

In recent years, large numbers of chronic psychiatric patients, many of them chronic schizophrenic and geriatric patients with organic brain syndromes, have been discharged from state hospitals and placed in communities that have no facilities to deal with them. The affected communities are up in arms not primarily because they are mental patients labeled with psychiatric diagnoses (because the majority are not recognized as expatients) but

because the *behavior* of some of them is sometimes imcomprehensible, deviant, strange, and annoying. In a similar fashion, in a study of psychiatric labeling and the rehabilitation of former psychiatric inpatients, Schwartz and colleagues[49] found that an expatient's level of impairment (his behavior) was far more important in determining whether or not he was rejected by the community than knowledge that the individual was receiving psychiatric treatment (and therefore was labeled mentally ill).

Rosenhan never considers the possibility that the negative attitude toward patients with psychiatric diagnostic labels might at least have *something* to do with the attitude of people toward the very behaviors that might be the basis for the diagnostic labels. For example, he says:

> The stigmatizing effects of psychiatric labels are so well known empirically and experientially (how might you feel if your colleagues believed you were a paranoid schizophrenic?) that it is hard to understand how or why those effects could be denied.[50(p1647)]

Does Rosenhan think the answer to his hypothetical question would be any different if put solely in behavioral terms without a diagnostic label—"how might you feel if your colleagues believed that you had an unshakeable but utterly false conviction that everybody was out to harm you"?

It is informative to consult the references[51-53] that Rosenhan[1] cited as offering data that "a psychiatric diagnosis is harmful." My interpretation of these data is that they merely show that the general public in a variety of ways ascribes a negative valuation to behavior that has been identified as mental illness. It is hard to see how the public would have a more positive attitude toward individuals with behavioral diagnoses unless you could convince the public that what was wrong with these individuals had nothing to do with mental illness. Merely changing the name of the type of mental illness will not eliminate the negative attitude. That is why the attempt every few decades to change the name of the condition given to individuals whose behavior is negatively evaluated by the public, with the hope of thereby changing the attitude towards such individuals, is largely doomed to fail. Recall how "psychopath" became "sociopath" and more recently "antisocial personality." Recall how the "sexual perversions" became "sexual deviations," which might become (according to a recommendation of the Task Force on Nomenclature and Statistics, Subcommittee on Sexual Disorders) "sexual object and situation disorders." As soon as everyone finds out what the new terms really mean, the basic attitude toward individuals with these conditions reappears.

Rosenhan does not propose the "mental illness as myth" notion, although why he does not is a mystery, since it clearly is consistent with his basic hypothesis that diagnoses are in the minds of the observers and not the behavior of the patients. Furthermore, the only way to avoid the stigma of the mental illness diagnoses that Rosenhan decries would be to do away with the concept of mental illness itself. Can this be done? In a fascinating study of psychiatric labeling among Eskimos, Jorubas, and other divergent groups, J. M. Murphy, PhD (unpublished data), noted the following:

> Explicit labels for insanity appear to exist in most groups. The labels refer to beliefs, feelings, and actions which are thought to emanate from the mind or inner state of an individual; they cause such persons to seek the aid of healers; and they bear strong resemblance to what we call schizophrenia. Of signal importance is the fact that the labels for insanity do not refer to one specific attribute but to a pattern of several interlinked phenomena. Despite wide variation in culture, a pattern composed of hallucinations, delusions, disorientations, and behavioral aberrations appears to identify the idea of "losing one's mind" almost everywhere even though the content of these behaviors is colored by cultural beliefs.

The implication is clear—mental illness is a label for phenomena that apparently exist in all cultures. Efforts to avoid the negative attitudes toward the phenomena by eliminating the label are misdirected. The most effective way of improving attitudes toward mental illness (as toward "cancer" or any other frightening illness) is to develop treatments that work and then convey this information to the public.

THE USES OF DIAGNOSIS

Rosenhan believes that the pseudopatients should have been diagnosed as having hallucinations of unknown origin. It is not clear what he thinks the diagnosis should have been if the pseudopatients had been sufficiently trained to talk, at times, incoherently, and had complained of difficulty in thinking clearly, lack of emotion, and that their thoughts were being broadcast so that strangers knew what they were thinking. Is Rosenhan perhaps suggesting multiple diagnoses of hallucinations, difficulty in thinking clearly, lack of emotion, and incoherent speech, all of unknown origin?

It is no secret that we lack a full understanding of such conditions as schizophrenia and manic-depressive illness, but are we quite as ignorant as Rosenhan would have us believe? Do we not know, for example, that hallucinations, in the context just described, are symptomatic of a different condition than are hallucinations of voices accusing the patient of sin, when associated with depressed affect, diurnal mood variation, loss of appetite, and insomnia? What about hallucinations of God's voice issuing commandments, associated with euphoric affect, psychomotor excitement, and accelerated and disconnected speech? Is this not also an entirely different condition?

There is a *purpose* to psychiatric diagnosis.[54] It enables mental health professionals to *communicate* with each other about the subject matter of their concern, *comprehend* the pathological processes involved in psychiatric illness, and *control* psychiatric disorders. Control consists of the ability to predict outcome, prevent the disorder from developing, and treat it once it has developed. Any serious discussion of the validity of psychiatric diagnosis or suggestions for alternative systems of classifying psychological disturbance must address itself to these purposes of psychiatric diagnosis.

In terms of its ability to accomplish these purposes, I would say that psychiatric diagnosis is moderately effective as a shorthand way of communicating the presence of constellations of signs and symptoms that tend to cluster

together and is woefully inadequate in helping us understand the pathological processes of psychiatric disorders; however, it does offer considerable help in the control of many mental disorders. Control is possible because psychiatric diagnosis often yields information of value in predicting the likely course of illness (eg, an early recovery, chronicity, or recurrent episodes) and because for many mental disorders (particularly the severe ones), it is useful in suggesting the best available treatment.

Let us return to the three different clinical conditions that I described, each of which had auditory hallucinations as one of its manifestations. The reader with any familiarity with psychopathology will have no difficulty in identifying the three hypothetical conditions as schizophrenia, psychotic depression, and mania. Anyone familiar with the literature on psychiatric treatment will know that there are numerous well-controlled studies[55] indicating the superiority of the major tranquilizers for the treatment of schizophrenia, electroconvulsive therapy for the treatment of recurrent unipolar depression and, more recently, lithium carbonate for the treatment of mania. Furthermore, there is convincing evidence that these three conditions, each of which is often accompanied by hallucinations, are influenced by separate genetic factors. As Kety[3] said, "If schizophrenia is a myth, it is a myth with a strong genetic component."

Should psychiatric diagnosis be abandoned for a purely descriptive system that focuses on simple phenotypic behaviors *before* it has been demonstrated that such an approach is more useful as a guide to successful treatment or for understanding the role of genetic factors? I think not. It is of interest that examination of the behavior therapy literature, which is full of theoretical attacks on the usefulness of psychiatric diagnosis, does not indicate that it has been abandoned in actual practice by behaviorally oriented therapists. The traditional diagnostic categories of anxiety neurosis, phobia, anorexia nervosa, obsessive-compulsive disorder, schizophrenia, and depression and sexual dysfunction, to name but a few, are apparently alive and well, and presumably responding to specific behaviorally oriented therapies. (I have a vision. Traditional psychiatric diagnosis *has* long been forgotten. At a conference on behavioral classification, a keen research investigator proposes that the category "hallucinations of unknown cause" be subdivided into three different groups based on associated symptoms. The first group is characterized by depressed affect, diurnal mood variation, and so on, the second group by euphoric mood, psychomotor excitement, etc.

If psychiatric diagnosis is not quite as bad as Rosenhan would have us believe, that does not mean that it is all that good. What *is* the reliability of psychiatric diagnosis? A review of the major studies of the reliability of psychiatric diagnosis prior to 1972, revealed:

> Reliability appears to be only satisfactory for three categories: mental deficiency, organic brain syndrome . . . and alcoholism. The level of reliability is no better than fair for psychosis and schizophrenia and is poor for the remaining categories.[56]

So be it. But where did Rosenhan get the idea that psychiatry is the only medical specialty that is plagued by inaccurate diagnosis? Studies have shown serious unreliability in the diagnosis of pulmonary disorders,[57] in the interpretation of electrocardiograms,[58] in the interpretation of roentgenograms,[59, 60] and in the certification of causes of death.[61] A review of diagnostic unreliability in other branches of physical medicine is given by Garland[62] and the problem of the vagueness of medical criteria for diagnosis is thoroughly discussed by Feinstein.[63] The poor reliability of medical diagnosis, even when assisted by objective laboratory tests, does not mean that medical diagnosis is of no value. So it is with psychiatric diagnosis.

Recognition of the serious problems of the reliability of psychiatric diagnosis has resulted in a new approach to psychiatric diagnosis—the use of specific inclusion and exclusion criteria, as in contrast to the usually vague and ill-defined general descriptions found in the psychiatric literature and in the standard psychiatric glossary of the American Psychiatric Association. This approach was started by the St Louis group associated with the Department of Psychiatry of Washington University[64] and has been further developed by my co-workers and myself[65] as a set of criteria for a selected group of functional psychiatric disorders, called the Research Diagnostic Criteria (RDC). The Table shows the specific criteria for a diagnosis of

Diagnostic Criteria for Schizophrenia From the Research Diagnostic Criteria*

A.	At least two of the following are required for definite and one for probable diagnosis:
	1. thought broadcasting, insertion, or withdrawal (as defined in this manual)
	2. delusions of control, other bizarre delusions, or multiple delusions (as defined in this manual), of any duration as long as definitely present
	3. delusions other than presecutory or jealousy, lasting at least one week
	4. delusions of any type if accompanied by hallucinations of any type for at least one week
	5. auditory hallucinations in which either a voice keeps up a running commentary on the patient's behavior or thoughts as they occur, or two or more voices converse with each other
	6. nonaffective verbal hallucinations spoken to the subject (as defined in this manual)
	7. hallucinations of any type throughout the day for several days or intermittently for at least one month
	8. definite instances of formal thought disorder (as defined in this manual)
B.	A period of illness lasting at least two weeks
C.	At no time during the active period of illness being considered did the patient meet the criteria for either probable or definite manic or depressive syndrome (criteria A and B under Major Depressive or Manic Disorders) to such a degree that it was a prominent part of the illness

*For what it is worth, the pseudopatient would have been diagnosed as "probable" schizophrenia using these criteria because of A6. In an oral communication, Rosenhan said that when the pseudopatients were asked how frequently the hallucinations occurred, they said "I don't know." Therefore, criterion A7 is not met.

schizophrenia from the latest version of the RDC.

Reliability studies utilizing the RDC with case record material (from which all cues as to diagnosis and treatment were removed), as well as with live patients, indicate high reliability for all of the major categories and reliability coefficients generally higher than have ever been reported.[66] It is therefore clear that the reliability of psychiatric diagnosis can be greatly increased by the use of specific criteria. (The interjudge reliability [chance corrected agreement, kappa] for the diagnosis of schizophrenia, using an earlier version of the RDC criteria with 68 psychiatric inpatients at the New York State Psychiatric Institute, was .88, which is a thoroughly respectable level of reliability.) It is very likely that the next edition of the American Psychiatric Association's *Diagnostic and Statistical Manual* will contain similar specific criteria.

There are other problems with current psychiatric diagnosis. The recent controversy over whether or not homosexuality per se should be considered a mental disorder highlighted the lack of agreement within the psychiatric profession as to the definition of a mental disorder. (It is difficult to determine at twilight whether it is day or night, but we have no such difficulty at midnight or noon. So too, our difficulty in defining the precise border of mental disorder and nonmental disorder in no way indicates the lack of utility of the concepts involved.) To the extent that our profession defines mental disorder as any significant deviation from the "good life" or "optimal human functioning," we will needlessly label many individuals as ill who are in no distress, function reasonably well, and hurt no one. This is a utopian conception of mental health that subjects the profession to the accusation that the sole function of the concept of mental disorder is social control and the pejorative labeling of all forms of social deviance. It is for this reason that we have proposed a more circumscribed definition,[54] but the criteria for this definition now appear to me to have incorrectly omitted certain, but by no means all, forms of antisocial behavior.

There are serious problems of validity. Many of the traditional diagnostic categories, such as some of the subtypes of schizophrenia and of major affective illness, and several of the personality disorders, have not been demonstrated to be distinct conditions or to be useful for prognosis or treatment assignment. In addition, despite considerable evidence supporting the distinctness of such conditions as schizophrenia and manic-depressive illness, the boundaries separating these conditions from other conditions are certainly not clear. Finally, the categories of the traditional psychiatric nomenclature are of least value when applied to the large numbers of outpatients who are not seriously ill. This may be a result of our greater ease in classifying conditions, such as the organic mental disorders and the psychoses, where the manifestations of the illness are qualitatively different from normal functioning. (For example, hallucinations are not part of normal functioning.) In contrast, with the personality disorders, we are dealing with quantitative variations in the intensity and pervasiveness of ubiquitous traits. (For example, some degree of suspiciousness or histrionic behavior is part of normal functioning.)

ROSENHAN'S RESPONSE TO CRITIQUES OF HIS STUDY

In his recent symposium article, "The Contextual Nature of Psychiatric Diagnosis," Rosenhan[43] responded to the critiques offered by me and by the other participants. My impending exhaustion, and I suspect, that of most readers who have gotten this far with my article, suggests the need for a *limited* critique of Rosenhan's article, which is as fascinating for what it omits as for what it says.

Let us start with the former. The interested reader can decide for himself which of the specific critiques of Rosenhan's original study that I presented in the earlier part of this article are telling and therefore worthy of a response from Rosenhan, and which his article neglected to discuss. I believe they are numerous.

What is most significant is his omission of any discussion of the crucial issue of how diagnostic labels given to pseudopatients with unusual complaints are at all relevant to the problems of the psychiatric diagnosis of real patients. There is also no discussion of the criticism that the concept of schizophrenia does not exclude the possibility of recovery, so there was no reason why the admission diagnosis of schizophrenia should have been revised merely because of normal behavior while in the hospital.

There is a most remarkable response to my demonstration that the pseudopatients were given an unusual diagnosis at discharge, "in remission," thus proving that the diagnoses given to the pseudopatients were a function of their behavior and not of the setting. Rosenhan states:

> Spitzer (1975) points out that the designation "in remission" is exceedingly rare. It occurs in only a handful of cases in the hospitals he surveyed, and my own cursory investigations that were stimulated by his confirm these observations. His data are intrinsically interesting, as well as interesting for the meaning they have for this particular study. How shall they be understood?
>
> Once again we return to the influence of context on psychiatric perception. Consider two people who show no evidence of psychopathology. One is called sane and the other is called paranoid schizophrenic, in remission. Are both characterizations synonymous? Of course not. Would it matter to you if on one occasion you were designated normal, and on the other you were call psychotic, in remission, with both designations arising from the identical behavior? Of course it would matter. The perception of an asymptomatic status implies little by itself. It is the context in which that perception is embedded that tells the significant story.[43(p468)]

Amazing! Of course "in remission" is not the same as "normal," but neither are the behaviors that are the basis for such categorization. The individual labeled "in remission" has the same behavior as the individual labeled "normal" only for one period of observation (current examination). He has a different history. The individual who has recovered (partially or completely) from an episode of schizophrenia has a probability of recurrence that far exceeds the probability for individuals who have never had such an episode of illness.[67] There is also evidence that maintenance phenothiazine treatment is effective in decreasing the probability of recurrence of a schizophrenic episode.[55] Therefore, if I, or a member of my family, had in fact recently recovered from an episode of schizophrenic illness and were currently asymptomatic, I

would *prefer* a diagnosis of schizophrenia in remission to a diagnosis of "normal," since it would suggest that a particular kind of treatment might well be indicated. That, after all, is one of the purposes of diagnosis. The same argument would apply with even greater force to the justification of the category "in remission" for individuals who have had recurrent episodes of depressive or manic illness or both and who at a particular examination are asymptomatic. Without the concept of affective disorder in remission, how could one justify the use of lithium carbonate as a prophylactic agent?

The justification of the category "in remission" for certain psychiatric disorders that tend to be chronic and recurrent in no way minimizes the difficulty in providing specific guidelines as to when use of this category is appropriate. Just as it would make no sense to use this category after all episodes of psychiatric illness regardless of type or duration of asymptomatic status, so it makes no sense to argue on logical grounds that it should never be used.

Rosenhan apparently believes he has discovered a scandal regarding the process by which the official nomenclature of the American Psychiatric Association is developed and adopted. He states:

> Unlike most medical diagnoses, which can be validated in numerous ways, psychiatric diagnoses are maintained by consensus alone. This is not commonly known to either the consumer or the mental health profession. Spitzer and Wilson[54] clarify the matter:
>
> > In 1965 the American Psychiatric Association ... assigned its Committee on Nomenclature and Statistics ... the task of preparing for the APA a new diagnostic manual of disorders ... A draft of the new manual, DSM II, was circulated in 1967 to 120 psychiatrists known to have special interests in the area of diagnosis and was revised on the basis of their criticisms and suggestions. After further study it was adopted by the APA in 1967, and published and officially accepted throughout the country in 1968.[43(p464)]

Does Rosenhan believe that it is only the psychiatric portion of the medical nomenclature that is decided on by a committee? (I am haunted by visions. The scene is Geneva, Switzerland, the offices of the World Health Organization. A senior official is overheard talking to a new employee who has recently joined his unit. "In order to keep on schedule, this would be a good time for you to go up to the mountain and bring down the stones that have the latest revision of the ICD [International Classification of Diseases] for the ninth edition. I wonder if He will make many changes?" The junior official, demonstrating his ignorance, asks, "Does He give us the entire medical classification that way?" The senior official replies: "Of course not. We only get the nonpsychiatric part that way. In order for us to get the psychiatric classification we have to have a committee that gets together and votes on the changes. It's all very messy and thank God it's only a small part of the entire medical classification.")

Classifications are all man-made, and either some single person or a group of individuals decides what they shall be. When a committee to develop a psychiatric classification functions, it can act, like any committee, wisely or foolishly. A committee acts foolishly if it has no clear understanding of the purposes of the classification it is developing and if it merely perpetuated traditional nosologic distinctions that are based on theoretical assumptions not supported by data. A committee acts wisely if it understands the multiple purposes of psychiatric classification, if it makes a serious effort to consider the data that have accumulated regarding such issues as the internal consistency of the phenomenology, differential response to treatment, outcome, familial pattern and genetic loading, and the understanding of basic psychopathological and physiological processes. A committee acts wisely if it consults with experts in the various areas under consideration and if draft proposals of the classification are subjected to public scrutiny.

Let us return again to Rosenhan's statement that "Unlike most medical diagnoses, which can be validated in numerous ways, psychiatric diagnoses are maintained by consensus alone."[43(p464)] Here Rosenhan is confusing the validation of the medical examination with the validation of a medical diagnosis. The distinction is subtle but important for understanding the true differences between psychiatric and nonpsychiatric medical diagnoses.

The validity of a procedure or concept is determined by how useful it is for the particular purposes for which it is intended. Consider the medical examination. The purpose is to make a correct diagnosis. Let us assume that a physician, during the course of a medical examination, determines on the basis of the patterning and course of symptoms and the physical examination that the most likely diagnosis is diabetes. The use of a laboratory procedure such as a glucose tolerance test can then validate the *examination* that led to the diagnosis of diabetes. The laboratory procedure, however, does not in any way validate the diagnostic category of diabetes. The validity of diabetes as a diagnosis is a function of the ability of physicians to understand and treat patients who have medical problems that are categorized as either diabetes or not diabetes. This is a function of our understanding of the illness, its course and associated features, and the availability of specific treatments. As greater understanding of the pathophysiology of diabetes and as more effective methods for treating diabetes are developed, the validity of the diagnostic category of diabetes increases.

Now let us consider the psychiatric examination. Its purpose is to make the correct psychiatric diagnosis. It is true that except for the organic brain disorders, we have no laboratory procedures that can be used to validate the psychiatric *examination*. That does not mean, as Rosenhan suggests, that we have no procedures for validating psychiatric diagnoses other than consensus. The procedures for validating psychiatric diagnoses do not differ in principle from those used to validate nonpsychiatric medical diagnoses. They consist of studies that indicate the extent to which knowledge of *membership* in a given diagnostic category provides useful information, not already contained in the defining characteristics of the diagnostic category. Reference has already been made to studies indicating the specificity of various forms of somatic treatment and separate genetic factors for several of the major psychiatric diagnostic categories. These studies, as well as other studies dealing with some of the other

purposes of psychiatric diagnosis, *are* the procedures by which psychiatric diagnoses are validated.

ROSENHAN'S SUGGESTIONS FOR THE FUTURE

Rosenhan[43] concludes his response to the critiques of his study with a section entitled "The Future." It begins as follows: "It is natural to infer that what I have written here argues against categorization of all kinds. But that is not the case. I have been careful to direct attention to the present system of diagnosis, the DSM-II [*Diagnostic and Statistical Manual of Mental* Disorders[48]]."[43(p472)] Only DSM II and not the general utility of psychiatric diagnosis? What about the following statements in the original article[1]?

> Psychiatric diagnoses, in this view, are in the minds of the observers and are not valid summaries of characteristics displayed by the observed... It seems more useful... to limit our discussions to behaviors, the stimuli that provoke them, and their correlates... Rather than acknowledge that we are just embarking on understanding, we continue to label patients "schizophrenic," "manic-depressive," and "insane," as if in those words we had captured the essence of understanding.[43(pp251-254,257)]

What about the following statement in the same article, in which Rosenhan claims to be discussing the DSM II only?

> Indeed, at present, my own preference runs to omitting diagnoses entirely, for it is far better from a scientific and treatment point of view to acknowledge ignorance than to mystify it with diagnoses that are unreliable, overly broad, and pejoratively connotative.[43(p467)]

Rosenhan continues:

> Nothing that is said here is intended to deprive the researchers of his classificatory system. He cannot proceed without it, but as long as his diagnostic data remain in his file until they are fully validated, they can do patients and treatment no harm.[43(p473)]

What in the world is a "fully validated" diagnostic system? What is the clinician, who has to do the best he can with what information is currently available, to do as he waits for the appearance of the "fully validated" classification? Rosenhan ignores the historical fact that classification in medicine has always been preceded by clinicians using *imperfect* systems that have been improved on the basis of clinical and research experience.

(Hang on reader. We are almost finished.) Rosenhan again states:

> What might we require of new diagnostic systems before they are published and officially accepted?... We should ask that coefficients of agreements between diagnosticians in a variety of settings *commonly* reach or exceed .90. That figure, which is associated with a bit more than 80% of the variance in diagnosis, is a liberal one in terms of the possible consequences of misdiagnosis and the reversibility of the diagnoses. The full reasoning behind that figure takes us away from the central thrust of this paper, but interested readers can confirm it for themselves in Cronbach, Gleser, Harinder, and Nageswari [68], and Cronbach and Gleser [69].[43(p473)]

First of all, the coefficients of agreement, such as the kappa index to which Rosenhan previously referred, unlike product moment correlations, are already in units of the proportion of subject variance and do not need to be squared.[70,71] Thus, a kappa of .8 means that 80% of the variance is associated with true subject variability.

The more important error is Rosenhan's justification of an entirely arbitrary requirement of a given level of interrater agreement by citing a tradition in psychometrics, which makes an assumption that cannot be made in psychiatric diagnosis. Rosenhan's reference to the two excellent psychometric textbooks discusses the desirability of avoiding decisions when the likelihood of an error exceeds .05 or .10. What about situations when a decision cannot be avoided, which is the general rule when a patient is examined psychiatrically? Let us take the example of a decision regarding suicidal behavior (which admittedly is not a diagnosis but illustrates the issues well). Interrater agreement regarding suicidal potential is undoubtedly much below .8. Does that mean that a clinician should never make a management decision based on his best judgment? Obviously, he must—to avoid making a decision is itself a decision. The facts are that despite our difficulty in reliably making medical judgments regarding diagnostic categories in psychiatry and the rest of medicine, patients must be treated and that treatment must follow from the decision of the clinician as to what he thinks is wrong with the patient. And that is what a diagnosis is.

Finally, Rosenhan concludes:

> We should require that the proven untility of such a system exceed its liabilities for patients. Understand the issue. Syphilis and cancer both have negative social and emotional overtones. But the treatments that exist for them presumably exceed the personal liabilities associated with the diagnosis.[43(p472)]

The implications of this are staggering! Is Rosenhan suggesting that prior to the development of effective treatments for syphilis and cancer, he would have decried the use of these diagnostic labels? Should we eliminate the diagnoses of antisocial personality, drug abuse, and alcoholism until we have treatments for these conditions whose benefits exceed the potential liabilities associated with the diagnosis? How do we study the effectiveness of treatments for these conditions if we are enjoined from using the diagnostic categories until we have effective treatments for them?

I have not dealt at all with the myriad ways in which psychiatric diagnostic labels can be and are misused and hurt patients rather than help them. This is a problem requiring serious research that Rosenhan unfortunately does not help illuminate. However, whatever the solutions to that problem, the available evidence that psychiatric diagnostic labels are *inherently* harmful to patients is scant indeed. Their misuse is not sufficient reason to abandon their use; when properly used, they have been shown to be of considerable value.

This article is a revision of that which appeared in the 1975 volumes of The Journal of Abnormal Psychology. *The need for serious consideration of issues that can so mischievously mislead the serious agenda of a field is well met therein; Professor Rosenhan also there amplifies his view of both the data and the controversy. The space and attention given the original article, in part because of its sponsorship, evoked concern among those researchers and clinicians working seriously in the range of issues relevant to general psychiatry. Rather than launch more attention to the* Science *article and deprive our contributors of* ARCHIVES *space for original research and scholarly communications (that*

are scrutinized by stiff peer review), we have awaited an article suitable to the issue of science and sanity. We recommend most highly The Journal of Abnormal Psychology*'s symposium and especially the masterful analysis of philosophy of science issues by I. E. Farber* (J Abnorm Psychol *84:442-452, 589-620, 1975)—ED.*

Jean Endicott, PhD, Joseph Fleiss, PhD, Joseph Zubin, PhD, Janet Forman, MSW, Karen Greene, MA, and Rose Bender assisted with the preparation of this article.

This article is a revision of one that appeared in *The Journal of Abnormal Psychology* (84:442-452, 1975).

Nonproprietary Name and Trademarks of Drug

Lithium carbonate—*Eskalith, Lithane, Lithotabs.*

References

1. Rosenhan DL: On being sane in insane places. *Science* 179:250-258, 1973.
2. Dimond RE: Popular opinion is not empirical data. *Clin Soc Work J* 2:264:270, 1974.
3. Kety SS: From rationalization to reason. *Am J Psychiatry* 131:957-963, 1974.
4. Rifkin A, Klein DF, Quitkin F, et al: Sane: Insane. *JAMA* 224:1646-1647, 1973.
5. Klein DF, Rifkin A, Quitkin FM: Sane: Insane. *JAMA* 226:1569, 1973.
6. Fleischman PR, Israel JV, Burr WA, et al: Psychiatric Diagnosis. *Science* 180:356-365, 1973.
7. Pattison EM: Social criticism and scientific responsibility. *J Am Sci Affil* 26:110-114, 1974.
8. Rabichow HG, Pharis ME: Rosenhan was wrong: The staff was lousy. *Clin Soc Work J* 2:271-278, 1974.
9. Shectman F: On being misinformed by misleading arguments. *Bull Menninger Clin* 37:523-525, 1973.
10. Abrahamson D: Procedure re-examined. *Lancet* 1:1153-1155, 1974.
11. Arthur RJ: Social psychiatry: An overview. *Am J Psychiatry* 130:841-849, 1973.
12. Burdsal C, Greenberg G, Timpe R: The relationship of marihuana usage to personality and motivational factors. *J Psychol* 85:45-51, 1973.
13. Sane: Insane, editorial. *JAMA* 223:1272, 1973.
14. Insane: Sane, editorial. *JAMA* 223:1381, 1973.
15. Glaser FB: Medical ethnocentrism and the treatment of addiction. *Int J Offender Ther Comp Crimino* 18:13-27, 1974.
16. Hoekstra A: Concerning science and society. *Clin Soc Work J* 2:299-306, 1974.
17. Kane RA: Look to the record. *Soc Work* 19:412-419, 1974.
18. Levy CS: On sane social workers in insane places. *Clin Soc Work J* 2:257-263, 1974.
19. Mark VH: A psychosurgeon's case *for* psychosurgery. *Psychol Today* 8:28ff, 1974.
20. Morse N: Some problems with insane institutions. *Clin Soc Work J* 2:291-298, 1974.
21. Oran D: Judges and psychiatrists lock up too many people. *Psychol Today* 7:20ff, 1973.
22. Perry TL, Hansen S, Tischler B, et al: Unrecognized adult phenylketonuria. *N Engl J Med* 289:395-398, 1973.
23. Scheff TJ: The labelling theory of mental illness. *Am Sociol Rev* 39:444-452, 1974.
24. Walden T, Singer G, Thomet W: Students as clients: The other side of the desk. *Clin Soc Work J* 2:279-290, 1974.
25. Daves WF: *Textbook of General Psychology.* New York, Thomas Y Crowell Co, 1975.
26. Davison G, Neale JM: *Abnormal Psychology—An Experimental Clinical Approach.* New York, John Wiley & Sons Inc, 1974.
27. Haber R, Fried A: *An Introduction to Psychology.* New York, Holt Rinehart & Winston Inc, 1975.
28. Hilgard E, Atkinson R, Atkinson R: *Introduction to Psychology,* ed 6. New York, Harcourt Brace Jovanovich Inc, 1975.
29. Kleinmuntz B: *Essentials of Abnormal Psychology.* New York, Harper & Row Publishers Inc, 1974.
30. Krech D, Crutchfield R, Livson N: *Elements of Psychology,* ed 3. New York, Alfred A Knopf Inc, 1974.
31. London P: *Beginning Psychology.* Homewood, Ill, Dorsey Press Inc. 1975.
32. McMahon FB: *Psychology: The Hybrid Science,* ed 2. Englewood Cliffs, NJ, Prentice-Hall Inc, 1974.
33. Rubinstein J: *The Study of Psychology.* Guilford, Conn, Dushkin Publishing Group Inc, 1975.
34. Ullman L, Krasner L: *A Psychological Approach to Abnormal Behavior,* ed 2. Englewood Cliffs, NJ, Prentice-Hall Inc, 1975.
35. Wrightsman L, Sanford F: *Psychology: A Scientific Study of Human Behavior,* ed 4. Monterey, Cal, Brooks/Cole Publishing Co, 1975.
36. Zimbardo P, Ruch F: *Psychology and Life,* ed 9. Glenview, Ill, Scott Foresman & Co, 1975.
37. Brown R, Herrnstein R: *Psychology.* Boston, Little Brown & Co, 1975.
38. Kimble G, Garmezy N, Zigler E: *Principles of General Psychology,* ed 4. New York, Ronald Press Co, 1974.
39. Lindzey G, Hall C, Thompson R: *Psychology.* New York, Worth Publishers Inc, 1975.
40. Crown S: On being sane in insane places: A comment from England. *J Abnorm Psychol* 84:453-455, 1975.
41. Farber IE: Sane and insane: Constructions and misconstructions. *J Abnorm Psychol* 84:589-620, 1975.
42. Millon T: Reflections on Rosenhan's "On being sane in insane places." *J Abnorm Psychol* 84:456-461, 1975.
43. Rosenhan DL: The contextual nature of psychiatric diagnosis. *J Abnorm Psychol* 84:462-474, 1975.
44. Spitzer RL: On pseudoscience in science, logic in remission, and psychiatric diagnosis: A critique of Rosenhan's "On being sane in insane places." *J Abnorm Psychol* 84:442-452, 1975.
45. Weiner B: On being sane in insane places: A process (attributional) analysis and critique. *J Abnorm Psychol* 84:433-441, 1975.
46. Hunter FM: Psychiatric diagnosis. *Science* 180:361, 1973.
47. Rosenhan DL: Psychiatric diagnosis. *Science* 180:365-369, 1973.
48. *Diagnostic and Statistical Manual of Mental Disorders,* ed 2. Washington DC, American Psychiatric Association, 1968.
49. Schwartz CC, Myers JK, Astrachan BM: Psychiatric labeling and the rehabilitation of the mental patient: Implications of research findings for mental health policy. *Arch Gen Psychiatry* 31:329-334, 1974.
50. Rosenhan DL: *Sane: Insane. JAMA* 224:1646-1647, 1973.
51. Nunnally JC Jr: *Popular Conceptions of Mental Health.* New York, Holt Rinehart & Winston Inc, 1961.
52. Sarbin TR: On the futility of the proposition that some people be labelled "mentally ill." *J Consult Clin Psychol* 31:447, 1967.
53. Sarbin TR, Mancuso JC: Failure of a moral enterprise: Attitudes of the public toward mental illness. *J Consult Clin Psychol* 35:159-173, 1970.
54. Spitzer RL, Wilson PT: Nosology and the official psychiatric nomenclature, in Freedman A, Kaplan H, Sadock B (eds): *Comprehensive Textbook of Psychiatry.* Baltimore, Williams & Wilkins Co, 1975, pp 826-845.
55. Klein D, Davis J: *Diagnosis and Drug Treatment of Psychiatric Disorders.* Baltimore, Williams & Wilkins Co, 1969.
56. Spitzer RL, Fleiss JL: A reanalysis of the reliability of psychiatric diagnosis. *Br J Psychiatry* 125:341-347, 1974.
57. Fletcher A: Clinical judgment of pulmonary emphysema-an experimental study. *Proc R Soc Med* 45:577-584, 1952.
58. Davies LG: Observer variation in reports on electrocardiograms. *Br Heart J* 20:153-161, 1958.
59. Cochrane AL, Garland LH: Observer error in interpretation of chest films: International investigation. *Lancet* 2:505-509, 1952.
60. Yerushalmy J: Statistical problems in assessing methods of medical diagnosis, with special reference to x-ray techniques. *Public Health Rep* 62:1432-1449, 1947.
61. Markush RE, Schaaf WE, Seigel DG: The influence of the death certifier on the results of epidemiologic studies. *J Natl Med Assoc* 59:105-113, 1967.
62. Garland LH: The problem of observer error. *Bull NY Acad Med* 36:570-584, 1960.
63. Feinstein A: *Clinical Judgment.* Baltimore, Williams & Wilkins Co, 1967.
64. Feighner JP, Robins E, Guze SB, et al: Diagnostic criteria for use in psychiatric research. *Arch Gen Psychiatry* 26:57-63, 1972.
65. Spitzer RL, Endicott J, Robins E: *Research Diagnostic Criteria (RDC).* New York, Biometrics Research, New York State Department of Mental Hygiene, 1974.
66. Spitzer RL, Endicott J, Robins E, et al: Preliminary report of the reliability of Research Diagnostic Criteria applied to psychiatric case records, in Sudilofsky A, Beer B, Gershon S (eds): *Prediction in Psychopharmacology.* New York, Raven Press, 1975, pp 1-47.
67. Gunderson JG, Autry JH, Mosher LR, et al: Special report: Schizophrenia, 1974. *Schizophrenia Bull* 9:15-54, 1974.
68. Cronbach LJ, Gleser GC, Harinder N, et al: *The Dependability of Behavioral Measurements: Theory of Generalizability for Scores and Profiles.* New York, John Wiley & Sons Inc, 1972.
69. Cronbach LJ, Gleser GC: Interpretation of reliability and validity coefficients: Remarks on a paper by Lord. *J Educ Psychol* 50:230-237, 1959.
70. Fleiss JL, Cohen J: The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educ Psychol Meas* 33:613-619, 1973.
71. Fleiss JL: Measuring agreement between two judges on the presence or absence of a trait. *Biometrics* 31:651-659, 1975.