The Journal of Mind and Behavior
Spring 1993, Volume 14, Number 2
Pages 203–206
ISSN 0271–0137

The Selling of DSM: The Rhetoric of Science in Psychiatry. Stuart A. Kirk and Herb Kutchins. New York: Aldine de Gruyter, 1992, 270 pages, $19.95 paper.

Reviewed by Donald M. Hayes, Sam Houston State University

Kirk and Kutchins begin with the fact that the primary justification for the development and adoption of DSM-III was that it would increase diagnostic reliability. In various contexts, they discuss the sources of diagnostic "error." One is *information variance*, the consequence of different clinicians asking different questions, receiving slightly different responses to the same questions, interpreting responses differently, and eliciting different behavior from the client. Another is *criterion variance*, the consequence of different opinions among clinicians about what information is relevant, how information should be interpreted, and what diagnostic category, if any, is appropriate. Other sources of "error" are differences in clinicians' past personal experiences, expectations, emotions, and illogical thinking, unfounded inferences, selective attention, and stereotypes.

Diagnostic reliability is also diminished by more or less deliberate decisions of clinicians based on such considerations as *regulating client flow* (using more rigid definitions of diagnostic criteria when the agency has a full case load, and more liberal definitions when in need of clients); *protecting clients from harm* (using the least serious appropriate diagnosis, to protect the client from the effects of labelling); *acquiring fiscal resources* (using a more serious appropriate diagnosis if necessary to help the client, or the agency, qualify for insurance coverage); *rationalizing decision-making* (finding a treatment which works, and then making the diagnosis which justifies that treatment); and *advancing a political agenda* (a narrow definition of mental disorder reduces the number of human troubles that can be treated; some diagnoses are offensive to certain populations, such as homosexuals; some diagnoses may provide a legal defense of criminal acts).

These errors stem from the clinician's discretion, which obviously should not be abused. It should be just as obvious that almost all of these sources of "error" are also beneficial to clients. Why, then, the urgency to eliminate them in order to increase diagnostic reliability? Why was it not concluded that

> knowledge about human problems, about the interaction of the mind and body and about the person and the environment were too elementary, too unrefined to fit into a rigid classification of discrete disorders . . . that knowledge of effective treatment was much too loosely tied to diagnostic categories to provide much help for treatment

---

Requests for reprints should be sent to Donald M. Hayes, Ph.D., Department of Sociology, Sam Houston State University, Huntsville, Texas 77341.

> planning . . . [that] diagnostic imprecision was neither troubling nor particularly harmful to therapeutic activity at its current stage of development. (p. 223)

As Kirk and Kutchins say, "The reliability problem could have been discarded as trivial" (p. 223).

The fact that clinicians differ in the questions they ask and the interpretations they make can be viewed as a strength. It is part of the reason that one clinician can succeed where another has not. Why should all be encouraged to ask the same questions, and to arrive at the same interpretations, just so diagnostic reliability may be increased? And why would a professional organization made up mostly of clinicians (The American Psychiatric Association) adopt a manual (DSM-III) that attempts to eliminate the strengths of clinical discretion through the use of structured interview schedules and rigid criteria of diagnosis? And has DSM-III eliminated most of these sources of "error," and has it increased diagnostic reliability? *The Selling of DSM* provides an answer to these questions.

Five facts stand out as a backdrop to this story. The first is that the treatment of [what are perceived as] personal problems was rapidly expanding. Second, psychiatrists had been reduced to a minority of mental health professionals (they are now less than 20% of all mental health professionals—the other 80% are psychologists, social workers, nurses, and marriage and family counselors). Third, psychiatry had become marginal within the medical profession. Fourth, the scientific validity of the core concept of the discipline, mental illness, was under attack by both insiders and outsiders. Fifth, the low level of diagnostic agreement (reliability) among psychiatrists had become symbolic of the profession's vulnerability to public and scientific criticism, and a major source of embarrassment to the profession. The makers of DSM-III focused on solving that problem.

One of the ironies of the story is the role played by biopsychiatry. Biopsychiatry "involves the search for physiological, genetic, and chemical bases for mental disorder and the development and use of psychopharmacological agents for treatment" (p. 10). According to Kirk and Kutchins, "Biopsychiatry is an attempt to secure a more powerful base for psychiatry within the jurisdiction of both medicine and mental health" (p. 10). And one way to increase power in the field of mental health was to take charge of the official nomenclature and criteria of diagnosis of mental illness, and so DSM-III was a major agent of this revolutionary change, moving psychiatry away from an etiological psychodynamic perspective, and toward a descriptive, "atheoretical" nosology.

How this revolution occurred has not been previously explored, and this is the main issue which Kirk and Kutchins address. They begin with "the most salient scientific problem—diagnostic reliability—and [trace] how the developers of DSM-III created, managed, and used this problem to reform the official diagnostic manual" (p. 13). The approach is sociological, emphasizing the mobilization of others by claimsmakers within the context of certain situational contingencies.

Diagnostic reliability was not a problem for most mental health workers. In fact, the earliest classification in the U.S. was developed by the federal government, and classification was, and is today, motivated by administrative and governmental needs, and not by demands from practitioners. Furthermore, the issue of diagnostic reliability diverts attention from the more critical problem of validity—two or more diagnosticians may be in perfect agreement about a diagnosis, and both (or all) of them may be wrong.

One of the early steps in the revolution was to transform the problem of reliability from one which everyone understood to one about which only experts were

qualified to write and speak. This was achieved in part by a paper, by Robert Spitzer and others, which introduced the kappa statistic; complicated the discussion with the concept of weighted kappa; tied both to a Fortran IV computer program that was inaccessible to many; and provided no criteria for interpreting kappa.

Then two teams of researchers, one at Washington University in St. Louis, the other at the New York State Psychiatric Institute affiliated with Columbia University, "set out to improve reliability and thereby protect the conceptual integrity of psychiatry by tinkering with the diagnostic discretion of practitioners and then measuring their success with kappa" (p. 48). (Kirk and Kutchins do acknowledge that early publications of these groups were directed toward researchers, and not toward clinical psychiatrists.) These two teams determined to eliminate information variance by structuring the psychiatric interview, and to eliminate criterion variance by defining diagnostic categories in terms of behavior. The head of the New York group was Robert Spitzer, who is the chief villain in this story. The need for these changes was advanced by research which exaggerated the problem of reliability in the past, and exaggerated the improvements made by these new innovations.

Meanwhile, work on DSM-III was underway. Robert Spitzer was appointed head of the DSM-III Task Force, in part due to his role in resolving the controversy over homosexuality as a mental disorder. On the basis of the research by the New York and St. Louis groups, the DSM-III Task Force "devised an 'atheoretical' diagnostic system that emphasized specific descriptions of directly observable behaviors" (p. 77). Kirk and Kutchins note that this "was a radical departure for American psychiatry, which had accepted as a first axiom that its task was to identify and treat causes, not symptoms" (p. 77). Those appointed to work with Spitzer on the DSM-III Task Force included another from the New York group, two from the St. Louis group, and another who "also had a long-term interest in a behavioral approach to diagnosis" (p. 98). This small group of five, homogeneous as it was, made "the essential decisions about [DSM-III's] approach, structure, and contents" (p. 99). The first draft, produced within the year, "incorporated all of the major innovations that were eventually included in DSM-III" (p. 99). This Task Force had expanded to 15 psychiatrists and four consultants by the time its work was completed in 1980, and hundreds had participated on various committees related to DSM-III, but Spitzer "did his best to discourage participation by antagonistic psychiatrists, not only on the main DSM-III Task Force but also on the many subcommittees. . ." (p. 100). Spitzer also "dismissed questions about the adequacy of representation by minorities and women" (p. 101) and when a committee of black psychiatrists asked for representation on the Task Force Spitzer suggested, as Kirk and Kutchins phrase it, "that there was no black psychiatrist among the 18 thousand members of the APA with adequate expertise to meet his criteria for membership on the DSM-III Task Force" (p. 102).

The main point of all this is that, although DSM-III was promoted as a scientific accomplishment, it was much less a product of science than of politics. Just as the change pertaining to homosexuality was political (the result of a vote, following intense lobbying by homosexuals), so were changes to Gender Identity Disorder of Childhood, Sexual Sadism, Sexual Assault Disorder, Self-defeating Personality Disorder, and the addition of Post-traumatic Stress Disorder. So was the virtual exclusion of those with a psychodynamic perspective, a majority of the APA membership, from participation in the process. So was the omission of the definition of mental disorders "as a subset of medical disorders," after strong objection by the

American Psychological Association. Theodore Blau, president of the American Psychological Association at the time [1977], in a letter to Jack Weinberg, the APA president, wrote: "Candidly DSM-III . . . is more of a political position paper for the American Psychiatric Association than a scientifically-based classification system" (p. 115).

The field testing of DSM-III, like the participation of hundreds on the various committees and task forces, helped secure its adoption. Large numbers of (volunteer) participants served to "legitimate" the claims of superior reliability, when in fact there was little if any improvement over levels of reliability that had previously been denigrated in the course of creating demand for DSM-III. But like the earlier studies of reliability which had used the new interview schedules and diagnostic criteria, the DSM trials also received the best possible interpretation. This included a flexible interpretation of the kappa statistic, made easier by its earlier obfuscation. According to Kirk and Kutchins,

> The developers of DSM-III repeatedly asserted that careful, systematic field trials established the improved scientific reliability of DSM-III. However, important information about the models and findings of the trials was never reported. . . . The field trials themselves could more accurately be described as uncontrolled, nonrandom surveys in which several hundred self-selected and unsupervised pairs of clinicians throughout the country attempt to diagnose nonrandomly selected patients and, after some sharing of information, made "independent" assessments of these patients. (p. 157)

There were six primary sources concerning reliability of DSM-III, and "Spitzer and his close associates were the primary authors of all of them" (p. 165). Furthermore, "Spitzer was both the prime mover in the development of DSM-III and the principal investigator in the National Institute of Mental Health-funded field trials" (p. 186). Kirk and Kutchins note the absence of an independent testing agency, such as the FDA for new pharmaceutical drugs.

The authors point out that DSM-III-R was published before it could benefit from research on DSM-III, and that DSM-IV is anticipated before it can benefit from research on DSM-III-R. They suggest that part of the reason for such rapid change is that it deflects criticism—why criticize a manual that is already being revised? They note that "within only 14 years (1979–1993) . . . four [manuals] will have been operative" (pp. 213–214). None of the revisions has been demanded by clinical practitioners, and each of them starts with "official questioning of the scientific status of the current nosology" (p. 214).

The book is very well written, and even the discussions of statistical and methodological problems are easily comprehensible. It deserves a prominent place in this continuing controversy.