Section 1

Present, past, and future
Beyond the DSM and ICD: a rationale for understanding and using descriptive psychopathology

The straight jacket imposed on psychiatry by the introduction of [DSM-III] and its successors, reverberating in Europe with the [IDC-10] has had a profound effect on the practice of psychiatry. An earlier generation’s more elegant constructions of a hierarchal basis for diagnostic classifications has disappeared, so that a patient now may end up with 3, 4, or even more DSM-IV diagnoses, and patients failing to fulfill one of the criteria for entry for a condition may be deemed not to be suffering from that condition, although logical clinical evaluation would suggest otherwise. Furthermore, many of the diagnostic inclusions are broad, ambiguous, and open to misuse in inappropriate settings. It seems the DSM-IV and its forerunners were created by committees which appear not to have been appreciative of the broad spectrum of conditions met with in psychiatric practice, and particularly in neuropsychiatry.1

Present-day psychiatry is dependent upon the Diagnostic and Statistical Manual for Mental Disorders (DSM) and International Classification of Diseases (ICD) classifications. The DSM has become the main reference source of recognized psychopathology and is the standard system for research subject selection worldwide. The ICD, rarely used in the USA, is favored in many parts of the world as a clinically useful document.2 The ICD offers separate research and clinical versions in several languages.3 Training programs in many parts of the world also rely on the DSM and ICD. Endorsed treatment algorithms are linked to classification labels. Manual category numbering is required for clinical documentation, and insurance reimbursement, and is referenced in legislative and legal proceedings. This dependence is accepted under the assumption that the manuals maximize reliability and contain validated conditions and groupings that encourage the best diagnostic decisions and treatment choices.

The dependence on the classification manuals has permitted a paradigm shift in psychiatry, particularly in the USA. The more leisurely psychological approach to patient care has been largely replaced by a primary care treatment model. Rapid diagnosis, followed by reflexive pharmacotherapy is encouraged.

Developed from an article, Vaidya and Taylor (2006).
to accommodate high patient turnover. “I don’t have enough time to see my patients . . . they only give me a half hour for intakes and follow-up visits” has become the mantra of psychiatry house officers.4

Applying the primary care model is facilitated by the ICD short descriptive prose or the DSM telegraphic lists. While the two systems differ in some categories (e.g. psychotic disorders, dementia, disorders in children and adolescents, and generalized anxiety disorder) and terminology (e.g. the ICD “organic” versus the DSM “secondary to” for syndromes with established etiology), both offer a skeletal view of psychopathology designed to be applied quickly.5 The time-consuming detailed investigation of the unfolding of the patient’s illness, the nuances of the sequence of symptom emergence, patterns of features, and the importance of some features over others is deemed superfluous and has been abandoned. Once diagnostic criteria are met, a treatment algorithm based on the DSM or ICD diagnosis is chosen. Treatment algorithms, often endorsed by an “expert” panel,6 can be applied as if cooking recipes.7

Paralleling the format changes has been an expansion in diagnostic choices from a handful of syndromes in DSM-II (APA, 1968) and the ICD-6 mental disorders section to presently over 280 options. The expansion is meant to assure recognition of any psychiatric affliction,8 implicitly promising that the classification contains all the known psychiatric conditions, that these conditions are sufficiently validated, and that the diagnostic criteria for each are reliable and sufficient to identify each condition. There should be no practical need to know more psychopathology than what is in the manuals. The promise, however, is unfulfilled, as validity is poor for many classification groupings (e.g. personality disorders, impulse control disorders) and the reliability of the systems is marginal.

The weakness in present classifications is illustrated in the startling and clearly implausible announcement that a study supported by the National Institutes of Mental Health in the USA determined that 55% of persons in the USA are at lifetime risk for psychiatric illness. In response, Paul McHugh, professor of psychiatry at Johns Hopkins and retired department chairman, blamed inexperienced interviewers relying on the DSM. He wrote:

In addition to relying solely on respondents’ yes or no answers to a checklist, the investigators are committed to employing the official Diagnostic and Statistical Manual of Mental Disorders – Fourth Edition (abbreviated DSM-IV), which bases all psychiatric diagnoses on symptoms and their course, not on any fuller knowledge of the person. It is as if public health investigators studying the prevalence of pneumonia over time in the American population were satisfied to call every instance of a cough with a fever and a mucoid sputum a case of pneumonia.9
Training of descriptive psychopathology relies on classification manuals

The primary care model has elicited a metamorphosis in psychiatric instruction, notably in a reduction in the teaching of the mental status examination and descriptive psychopathology. Once the lynchpin of training, interest in psychopathology now focuses on how to recognize the clinical features needed to apply DSM or ICD labels. For example, in a 2002 mailed survey to all accredited psychiatry residency training programs in the USA (N=149), of which 68 (45.6%) responded, while nearly 80% stated that they offered a course in descriptive psychopathology (often only one semester), and another in the mental status examination (typically less than 5h), less than 30% of respondents taught the classic features of psychopathology (e.g. catatonia, first rank symptoms), and less than 20% used any of the well-known psychopathology texts. Twenty percent of programs offered no formal lecture series in descriptive psychopathology or mental status examination. Psychopathology was seen in many teaching programs as the signs and symptoms described in the present DSM, but nearly half did not provide classroom instruction or discussion of the features in the criteria, and those that did typically devoted less than 5h to it. A 1991 survey of all psychiatric clinical tutors in the UK also found substantial reliance on the DSM for the teaching of basic psychopathology. Surveys of the teaching of psychopathology in other parts of the world are lacking.

Problems in present classifications

Table 1.1 displays the problems in present classification. These are discussed in detail below.

DSM and ICD reliability is weak

Reliability in diagnosis is its degree of precision, i.e. agreement among clinicians. If reliability is poor, validity of diagnoses (accuracy) is unclear. Systematized “field trials” of the interrater reliability of the recent DSM and ICD iterations describe mixed results. Diagnostic agreement was also inflated by defining agreement as two clinicians placing the patient in the same diagnostic class rather than explicitly agreeing. If the clinicians differed in the specific disorder within the class (e.g. if one diagnosed “schizophrenia”, while the other said the patient had “delusional disorder”), agreement was accepted. Such agreement is equivalent to clinicians agreeing that a patient has a respiratory problem, but not whether it is bronchitis or pneumonia, bacterial or viral or allergy-related.
### Table 1.1. Problems in present classification

<table>
<thead>
<tr>
<th>Problem</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reliability is weak</td>
<td>The “claim to fame” of recent DSM iterations is high reliability. Weak reliability insures idiosyncratic diagnosis.</td>
</tr>
<tr>
<td>Encourages false positives and false negatives with over-inclusive diagnostic criteria</td>
<td>Some conditions are over-diagnosed (e.g. depression) resulting in false positives, research sample heterogeneity, and unneeded or dangerous treatment for patients. Some conditions are not recognized (e.g. catatonia), resulting in false negatives and inappropriate treatment.</td>
</tr>
<tr>
<td>Offers false choices</td>
<td>False choices lead to the prescription of inappropriate treatments. Conditions such as schizophreniform have no validity. Dissociation, a symptom, is treated as a disease. Abnormal bereavement and puerperal depressions are given separate status.</td>
</tr>
<tr>
<td>Omits or marginalizes established syndromes</td>
<td>Catatonia is incorrectly linked to psychotic disorders, melancholia is reduced to a modifying term, the different frontal lobe syndromes are not included.</td>
</tr>
<tr>
<td>Diagnostic criteria are poorly defined</td>
<td>Terms such as “disorganized speech” encourage misdiagnosis (e.g. misidentifying a fluent aphasia as flight-of-ideas or formal thought disorder).</td>
</tr>
<tr>
<td>Checklist format limits meaningful examination</td>
<td>Items are incorrectly given equal weight. Type of illness onset, sequence of symptom emergence, and patterns of features are mostly ignored, resulting in misdiagnosis.</td>
</tr>
<tr>
<td>Omits important discriminating psychopathology</td>
<td>Psychopathology associated with neurologic syndromes (e.g. psychosexual features and seizure disorder) are not mentioned, resulting in illnesses going unrecognized.</td>
</tr>
<tr>
<td>Claiming theory neutrality, it avoids neuroscience and laboratory criteria</td>
<td>Patterns of features that indicate the involvement of a specific brain region or system are not included (e.g. features indicating right hemisphere disease). Laboratory assessments are not included as helpful criteria (e.g. hypothalamic pituitary functioning in depressive illness, CPK levels and response to lorazepam in catatonia).</td>
</tr>
<tr>
<td>Longitudinal criteria are not used</td>
<td>The pre-psychosis findings in schizophrenia are ignored, resulting in the over-diagnosis of the condition and research sample heterogeneity. The dimensional traits of personality are ignored, resulting in poor reliability and validity for the personality disorders.</td>
</tr>
</tbody>
</table>
In the two DSM-III trials, the overall agreement for Axis I for adults was marginally acceptable (kappas of 0.68 and 0.72, with 0.70 the minimal agreement coefficient). For affective disorders they were 0.69 and 0.83. For schizophrenia, both trials obtained kappas of 0.81. The range among diagnostic classes, however, was broad and many had kappas near 0.50 (i.e. closer to chance agreement).\(^\text{18}\) Kappas for children and adolescents were poor for most conditions, as were kappas for Axis II. Often, only a few patient vignettes per category were used, lessening the likelihood of meaningful agreement. DSM-III-R field trials have similar shortcomings (APA, 1987).

Assessments of DSM-IV also detail mixed results. The DSM-based clinical interview is reported to reliably identify patients with eating disorders,\(^\text{19}\) symptoms and diagnoses in relatives of psychiatric patients,\(^\text{20}\) and diagnoses from information obtained from personal interview or from an informant.\(^\text{21}\) Other reports are less positive. For example, an assessment of 362 outpatients using the DSM-IV interview instrument for anxiety and mood disorder obtained good test–retest reliability for the two categories, but there was substantial overlap and “a common source of unreliability was disagreements on whether constituent symptoms were sufficient in number, severity, or duration to meet DSM-IV diagnostic criteria.”\(^\text{22}\) Participants also had difficulty categorizing clinical features presented randomly as representing an Axis I or II criterion, and in one study they misclassified 31% of Axis I criteria as representing an Axis II disorder and 25% of Axis II criteria as representing an Axis I condition.\(^\text{23}\) They could not identify whether a feature was a symptom of disease or trait behavior. This is equivalent to not knowing whether the patient’s cough represents a nervous tic or respiratory disease.

The DSM-IV field trials revealed diagnostic uncertainty. The mood disorder field trials of 524 patients from inpatient, outpatient, and community settings from 5 sites used structured interviews and reported good intra-site but only fair inter-site reliability with deteriorating reliability in a six-month retest.\(^\text{24}\) Test–retest reliability is reported below for statistical reliability standards for psychosis, somatization, eating disorder, dysthymia, mania, generalized anxiety disorder, attention deficit hyperactivity disorder, and hypochondriasis.\(^\text{25}\) Studies of the multiaxial systems of both classifications find poor agreement (i.e. which axis to place clinical features) and poor agreement on axes assessing environmental stressful events.\(^\text{26}\)

The reliability of “bizarre” delusions, the hallmark of the criteria for the psychotic disorders, has also been found unsatisfactory.\(^\text{27}\) An assessment for substance-induced psychiatric syndromes in 1951 acute psychiatric inpatients found a dimensional approach to have better reliability and predictive validity than the dichotomous DSM-IV strategy.\(^\text{28}\)
ICD-10 field trials worldwide assessed over 15,000 patients at 112 clinical centers in 39 countries. Good reliability was achieved except for the personality disorders.\textsuperscript{29} Independent examinations of 150 patients assessed with a European diagnostic instrument also found good reliability for schizophrenia,\textsuperscript{30} mania, and major depression, but unsatisfactory reliability for schizoaffective disorder.\textsuperscript{31} The validity of the psychotic disorders category, however, was questioned.\textsuperscript{32} Systematic application of ICD descriptions to clinical samples also finds instability over time for the diagnosis of bipolar and recurrent depressive disorder.\textsuperscript{33} The low interrater reliability for a depressive episode\textsuperscript{34} and difficulties with the ICD depression subscales for endogenous and psychogenic depression\textsuperscript{35} partially account for the diagnostic instability of the ICD mood disorder category. ICD reliability was found enhanced by the addition of clinical descriptions to the operational criteria, a strategy not used in the DSM.\textsuperscript{36}

The mixed reliability results are particularly alarming because the field trial participants were intensively trained in the use of the system and examination instruments. Also, about 40\% of the patient evaluations were done conjointly. These procedures are rarely used in clinical practice where diagnostic agreement among clinicians remains low.\textsuperscript{37}

Further, while the assessment of patients by structured examination can obtain fair to good reliability,\textsuperscript{38} this method has poor agreement with the more likely clinical circumstance of a clinician doing a semi-structured evaluation\textsuperscript{39} or a standard clinical assessment.\textsuperscript{40} Even when using semi-structured assessments, reliability is marginal for some diagnostic options.

Bertelsen (1999) cautions against the exclusive reliance on simplified list-based criteria. He points out that the best clinical approach is an initial comprehensive traditional clinical examination to first identify the syndrome followed by the matching of the findings to criteria for nosologic labeling, rather than reliance solely on the manuals.\textsuperscript{41}

\textbf{Classification validity is uncertain}

Accuracy in diagnosis defines validity, i.e. the patient has the illness that is diagnosed. Poor validity leads to false positive and false negative classifications. A false negative occurs when the patient’s illness is unrecognized. A false positive occurs when a patient is given a diagnosis he does not have. Present classification methods encourage both types of errors.

\textbf{Diagnostic false positives}

Diagnostic criteria in the manuals are mostly imprecise and overly broad, encouraging the identification of illness when none exists, or misidentifying one illness for another. The identification of over half the population in the USA as meeting
such criteria for illness dramatically demonstrates the degree of false positive
diagnosis inherent in using the DSM. Half of persons in normal bereavement also
meet criteria for major depression, but neither the researchers nor the bereaved in
the studies considered the state to be illness. The major depression criteria of
apathy and motor slowing are seen in frontal circuitry disease, while low energy,
shyness, and anxiety are found in some persons with personality deviations.
These patients may be misdiagnosed as depressed and needlessly prescribed
antidepressant agents.

The DSM diagnosis of major depression requires five or more items in any
combination. Depressed mood need not be present for the diagnosis of depres-
sion. A loss of interest or the inability to experience pleasure are acceptable
alternatives. “Fatigue or loss of energy” and “diminished ability to think or
concentrate” are choices. The criteria are not operationally defined (e.g. what
degree of diminished concentration is needed to be a symptom and how concen-
tration is to be measured, are not detailed). In the quest for diagnostic reliability,
criteria are over-simplified, thereby lowering the bar for admission into the
category of depression. Taken literally (which is a necessity to obtain expected
reliability) the following patient meets DSM-IV criteria for major depression.

Patient 1.1
A 51-year-old man experienced substantial loss of interest and anhedonia for
almost a year. He slept much of the day (hypersomnia is a criterion choice),
and his movements and thinking were slowed (psychomotor retardation is a
criterion choice). He had trouble concentrating his thoughts, and had no
energy. He was pessimistic about the future. He did not want to kill himself,
but he did not want to live in his present state. His symptoms caused
“clinically significant distress and impairment in social functioning.” His
condition could not be explained as the “direct physiological effects of a
substance . . . or a general medical condition.” His general neurologic exami-
ation was normal, except for slowness of movement and thought. His symptoms
began after his trailer home burned, destroying it and all his possessions. He
was not burned and did not suffer significant smoke inhalation. Posttraumatic
stress disorder was ruled out, because he did not have nightmares and was
neither anxious nor ruminating about the event. Major depression was diag-
nosed by several clinicians and antidepressant medications were prescribed
without improvement.

On examination, the man’s mood was reactive, and although subdued, he
showed mildly diminished emotional expression rather than sadness or appre-
hension. A frontal lobe avolitional syndrome was diagnosed and carbon
monoxide poisoning hypothesized as the cause of his behavioral change.
CT scan showed bilateral basal ganglia calcifications, a finding consistent with the diagnosis of carbon monoxide exposure. Methylphenidate treatment improved his condition.

Patient 1.1 also meets criteria for “treatment-resistant depression”, because he did not respond to two drug trials with different classes of antidepressants. However, about 10–15% of depressed patients labeled “treatment-resistant” are incorrectly considered depressed, and therefore do not benefit from antidepressant treatments.47

The DSM criterion A common to all the psychotic disorders is also problematic. Two of five features are needed, but sustained auditory hallucinations and “bizarre” delusions may stand alone. This provision is a vestige from the ideas of Kurt Schneider, who considered some psychotic features to be pathognomonic of schizophrenia if a neurological disease could not be recognized.48 The identification in the 1970s of Schneider’s “first rank symptoms” in patients with mood disorder and other conditions, however, demonstrated definitively that these features are not pathognomonic, but the error persists in DSM-IV. Consider Patient 1.2.

Patient 1.2

A 32-year-old woman was hospitalized because she barricaded her home and rearranged the furniture so that her two young children would not have to walk on the floor. She said she had overheard neighbors constantly plotting to electrify the floor and that she could feel static electricity. She was irritable and walked constantly throughout the inpatient unit on tiptoe and had several other catatonic features.

She responded to questions such as “What do you think is the reason for your neighbors doing those things to your house?” with:

“They’re jealous, mean spirited, I’m the spirit of 1776, they see the spirit in me, I have an aura, an aura borealis, a whore (eyes filled with tears for a moment), a four by four.”

Patient 1.2 meets the DSM criterion A for a psychotic disorder. She experienced sustained auditory hallucination (tactile also) and many would accept her delusional ideas as “bizarre”. She also exhibited “disorganized speech”, another criterion A choice. However, other psychopathology can be recognized. Her tiptoe gait is consistent with catatonia, and she exhibited other catatonic features, phenomena not detailed in the manuals.49 Criterion A includes catatonia as a choice in the diagnosis of schizophrenia. Nevertheless, irritability and constant walking about the inpatient unit suggests hyperactivity or agitation and along with catatonia are consistent with a manic episode. She had grandiose delusions. Characterizing speech with such vague terms as “disorganized” is also poor practice.
Flight-of-ideas with clang associations describes her language better, and are features of mania. The patient was treated with lithium monotherapy and fully recovered.

False negative diagnosis and “not otherwise specified” (NOS)

The high proportion of patients receiving the DSM Not Otherwise Specified (NOS) choice further attests to the limits of the system. To support treatment choices, the “catch-all” option permits clinicians to assign patients to a likely diagnostic category despite being unable to fit them to a specific illness descriptor (e.g. the diagnosis “psychosis, NOS” justifies prescribing an antipsychotic agent).

Use of the NOS choice occurs in several circumstances. Most commonly, the patient meets some but not all necessary criteria. When a patient has an established syndrome not recognized in the DSM, but has a clinical feature that suggests a diagnostic category, NOS is also applied. The frontal lobe avolitional and disinhibited syndromes, several seizure-related syndromes and the paraphrenias are not included in the DSM. Patients with these conditions go unrecognized and are typically labeled “psychotic disorder” or “mood disorder, NOS”.

Hirschfeld (2001) reviews the behaviors consistent with a manic-depressive spectrum, a construct not implicitly incorporated in the DSM. Such syndromes elicit the NOS suffix. The manic-depressive spectrum concept, however, leads to more effective treatment (e.g. mood stabilizers and antidepressants rather than psychotherapy alone) for many patients now considered as having personality disorders. Cyclothymia represents part of that spectrum.

The Oneiroid Syndrome, a dream-like state, known to European psychiatrists, but all but forgotten in the USA, is another example. Recent reviews of the diagnostic usefulness of psychopathology associated with traumatic brain injury and epilepsy further highlight the omission of important syndromes.

The failure to define the catatonia syndrome illustrates another ICD and DSM shortcoming. Catatonia has strong linkage to mood disorder, more so than to schizophrenia. Yet, the DSM primarily places catatonia as a subtype of schizophrenia, while all patients with catatonia not clearly the result of a neurologic or general medical condition must be diagnosed as suffering from a psychotic disorder by the ICD. There are over 40 classic catatonic features and associated behaviors, but the DSM briefly mentions only 12 and the ICD fewer. Neither manual offers instructions on how to identify or elicit the features. A patient could easily have many catatonic features not elicited or recognized by the clinician trained to the DSM or ICD standard. It is not surprising that most DSM-trained clinicians think catatonia is rare despite the consistent finding that when systematically assessed, 10% of acutely hospitalized psychiatric patients, 40% of hospitalized manic patients, and many patients with developmental