

Ethical Considerations in the Development and Application of Mental and Behavioral Nosologies: Lessons from DSM-5

Robert M. Gordon · Lisa Cosgrove

Received: 1 October 2013 / Accepted: 11 October 2013 / Published online: 5 December 2013
© Springer Science+Business Media New York 2013

Abstract We are not likely to find a diagnostic system as “unethical,” per se, but rather find that it creates ethical concerns in its formulation and application. There is an increased risk of misuse and misunderstanding of the DSM-5 particularly when applied to forensic assessment because of documented problems with reliability and validity. For example, when field tested, the American Psychiatric Association reported diagnostic category kappa levels as acceptable that were far below the standard level of acceptability. The DSM-5 does not offer sensitivity and specificity levels and thus psychologists must keep this in mind when using or teaching this manual. Also, especially in light of concerns about diagnostic inflation, we recommend that psychologists exercise caution when using the DSM-5 in forensic assessments, including civil and criminal cases. Alternatives to the DSM-5, such as the International Classification of Diseases and the Psychodynamic Diagnostic Manual are reviewed.

Keywords Diagnostic and statistic manual of mental disorders · DSM-5 · Ethics · Forensic assessment · International classification of disease · Mental and behavioral nosologies · Psychodynamic diagnostic manual

Problems of Development

A scientific discipline starts with a classification of its’ constructs. The developing taxonomy defines the constructs, their

validity, reliability, and inter-relationships for the purpose of understanding nature and making predictions. The particular taxonomy of diseases, i.e., nosology, used in a field carries the additional burden of utility, because it is an instrument of health care. Furthermore, taxonomy of an individual’s traits, behaviors, and symptoms has implications not found with other classification systems. A diagnosis of a mental or behavioral disorder can affect the type and dose of treatment. A missed or inaccurate diagnosis can delay treatment or offer the wrong treatment and make suffering worse. It can bring stigma, and affect finding or retaining employment. In the forensic arena, it can affect criminal sentencing or compensation from injury.

The latest revision to the Diagnostic and Statistical Manual of Mental Disorders (DSM)-5 (American Psychological Association 2013), includes a “Cautionary Statement for Forensic use of DSM-5” “...When used appropriately, diagnoses and diagnostic information can assist legal decision makers in their determinations... there is a risk that diagnostic information will be misused or misunderstood.” (p. 25). Indeed, there is an increased risk of misuse and misunderstanding of the DSM-5 when applied to forensic assessment because of documented problems with reliability and validity.

Pay Attention to the Statistics

Many of the diagnoses have very low inter-rater agreement (kappa levels) and there is no sensitivity or specificity data reported for the DSM-5 (Karson 2010). Regier et al. (2013) reported that kappa coefficients for some of the most common disorders such as generalized anxiety disorder are as low as 0.20. Concerns over reliability led the former chair of the third edition of the DSM, Robert Spitzer, (Spitzer, Williams, & Endicott, 2012, p. 537) to caution that, “...calling for psychiatry to accept kappa values that are characterized as unreliable in other fields of medicine is taking a step backwards.” Indeed, diagnostic reliability is essential for the credibility of a

L. Cosgrove
University of Massachusetts-Boston, Cambridge, USA

L. Cosgrove
Edmond J. Safra Center for Ethics, Harvard University,
Cambridge, USA

R. M. Gordon (✉)
Independent Practice, Allentown, PA, USA
e-mail: rmgordonphd@gmail.com

diagnostic manual, and kappa values of less than 0.60 are considered “cause for concern” (Spitzer et al. 2012). The former chair of the fourth edition of the DSM, Allen Frances, stated this point more bluntly: “DSM 5 has flunked its reliability tests” (Frances 2012). The published DSM-5 field trials reported kappa scores that showed poor test–retest reliability for many diagnoses (e.g., major depressive disorder: 0.32; mild neurocognitive disorder: 0.50; Brauser 2012; Frances 2012). Regier et al. (2013) found that of 15 adult and 8 child/adolescent diagnoses for which adequate sample sizes were obtained to report adequately precise estimates of the intraclass kappa, only five diagnoses were in the very good range (kappa = 0.60–0.79). Moreover, diagnostic reliability will most likely be even lower in real-world settings because these results were obtained under ideal conditions “in academic settings with trained and skilled interviewers, highly selected patients, and no time pressure” (Frances 2012). Yet despite the overall poor kappas for many DSM-5 diagnoses, the American Psychological Association (APA) declared: “Most diagnoses adequately tested had good to very good reliability” (Regier et al. 2013).

Arbitrary Cutoffs without Empirical Evidence

The DSM-5 often includes the use of non-scientifically determined thresholds. The DSM uses discrete time periods (i.e., for a period of 6 months) in criteria for arriving at diagnoses in order to increase reliability but this procedure might serve to decrease validity. It is easier to classify based on a distinct threshold date. For example, the DSM-5 diagnosis of pedophilic disorder requires, “... Over a period of at least 6 months...” (p. 697). The International Classification of Diseases (ICD-10), in contrast, offers the much more flexible standard, “A persistent or a predominant preference for sexual activity with a prepubescent child or children” p.166.

The DSM also employs a specific number of symptoms from categories to achieve a diagnosis. For example, an acute stress disorder diagnosis must have “nine (or more) of the following symptoms from any of the five categories...” p. 280. The specified number of symptoms required for a diagnosis is concrete but hardly scientific. Such specification may be helpful to researchers who need to agree on a precise definition of an independent variable. However, non-empirically based criteria complicate assessments for forensic psychologists by giving the impression of precision without validity.

Problems with an Atheoretical Symptom-Based Taxonomy

Since the DSM-5 is atheoretical and based mainly on a consensus about symptom clusters, the number of diagnoses can continue to expand or be relabeled without a sufficient

scientific basis. The DSM-5 introduced new or revised diagnoses such as binge eating disorder, autism spectrum disorder, disruptive mood dysregulation disorder in children, mild neurocognitive disorder, and premenstrual dysphoric disorder. In addition to the newly included diagnoses, perhaps one of the most controversial revisions in the DSM-5 is the elimination of the bereavement exclusion from the diagnostic criteria for a major depressive episode. With this change, individuals who are actively grieving a loss may be diagnosed with major depressive disorder (if they present with symptoms of depression 2 weeks after the loss). Some have argued that eliminating the bereavement exclusion for major depressive disorder would create a “false positive problem”—that people who are going through the normal process of grieving would be diagnosed with depression—which in turn, would create large opportunities for industry profit (Frances 2012). Indeed, the absence of biological markers for any of the DSM-5 disorders renders psychiatry more vulnerable to overdiagnosis and overtreatment of conditions that lack validity as distinct clinical entities.

Thus, it is not surprising that the revision process created a firestorm of controversy—mainly because of concerns about the widening of diagnostic boundaries. For example, shortly after the publication of the DSM-5 in May 2013, the director of the National Institute of Mental Health (NIMH), Dr. Thomas Insel, announced that NIMH would be “...re-orienting its research away from DSM categories” because they lack validity and “patients deserve better” (<http://www.nimh.nih.gov/about/director/2013/transforming-diagnosis.shtml>). However, NIMH’s proposed reductionistic taxonomy may only be useful for disorders that have more evidence of a neurobiological etiology (e.g., schizophrenia) and has limited clinical utility for disorders of relationships, trauma, and personality.

The DSM-5 favors behavioral, overt descriptions of disorders that help mental health practitioners to easily identify disorders with little training. But these behavioral symptom-based descriptions cannot provide a contextual understanding of mental disorders. Such an understanding must include subjective mental life, an appreciation for the unique life circumstances and lived experience of the client, and thus would require training into psychodynamics as well as the talent for insight. Research on personality indicates that it is complex and interactive and not merely a sum of its easily observable parts (Westen, Defife, Bradley, & Hilsenroth 2010; Young 2011).

Conflict of Interest

There is also the concern that industry and guild interests may have exerted undue influence over the revision process. Taxonomy of mental and behavioral disorders is not based on research alone, but evolves in the context of the needs and

beliefs of the sponsoring organization. The DSM was authored exclusively by the American Psychiatric Association, the manual brings considerable recognition, power, and money to the association, and thus, there are strong guild interests at stake. It is noteworthy that a paradigm shift occurred with the publication of the DSM III in 1980—the DSM III adopted a disease model and thus secured organized psychiatry's reputation as a *medical* specialty. This shift to a medical model elevated psychiatry's professional reputation. This elevation in status continues to give psychiatry—and thus the DSM—legitimacy, but it also opened the door to an improper dependence on the pharmaceutical industry (Cosgrove and Wheeler 2013). It is the promotion and dissemination of the DSM-5 as a scientifically valid and empirically based instrument that fosters acceptance of a disease model of mental disorders. This view is embraced by drug firms and believed by the general public. Although it was not the APA's or Robert Spitzer's (the chair of DSM III) intention to develop a psychiatric manual that was an industry friendly instrument, Dr. Spitzer later acknowledged that “[t]he pharmaceuticals were delighted” (Ronson 2011), with the model the DSM III adopted.

Although each subsequent DSM revision was described as being scientifically valid and evidence-based, it is important to remember that the DSM-5 was developed with an emphasis on consensus. As Karson (2010) noted, “The social or political importance of the DSM can be inferred from its charge to its work groups to form a consensus rather than to find out about nature...” (p. 2) A proposed diagnosis may be valid but orphaned with too few advocates to be voted as an official diagnosis or at the other extreme, accepted as an “official” diagnosis because it is popular or has economic value despite its questionable validity.

Alternatives to the DSM-5

In contrast to the DSM-5, the ICD—Mental and Behavioral Disorders is a product of the World Health Organization (WHO), the purpose of which is to classify and record the incidence of disorders worldwide. They seek input from several mental health professional groups from many different cultures. Evans et al. (2013) surveyed 2,155 psychologists from 23 countries, regarding diagnostic classification systems for mental and behavioral disorders in order to inform the development of the upcoming ICD-11 by the WHO. Psychologists thought that informing treatment decisions and facilitating communication were the most important purposes of classification, and preferred flexible diagnostic guidelines to strict criteria.

The newest psychological taxonomy is the Psychodynamic Diagnostic Manual (PDM Task Force 2006). The task force worked with five different psychoanalytic organizations to

form the taxonomy of the whole person. It is not a disease model but rather runs the descriptive range from very disturbed to healthy and puts personality structure at the heart of the taxonomy.

The PDM may, in some situations, be better at both informing treatment recommendations and risk management than the DSM or ICD. For example, the DSM classifies antisocial and borderline personality disorders by precise and narrow symptoms. This is often misleading. Psychopathy can be a complex personality pattern that combines with or is obscured by other personality patterns, and borderline can be viewed as an entire level of personality organization that can be applied to the various personality disorders. Since the PDM does view “borderline” as a level of severity of any personality disorder and psychopathy as a complex personality pattern, the practitioner may be less likely to fall into ethical conflicts by recognizing problematic behaviors early on (Gordon 2007). Also, prototypic criteria as used by the PDM offers a much more flexible, rich, and reliable description of diagnoses than the categorical criteria of the DSM-5 (Bornstein and Gordon 2012; Westen et al. 2010). The weakness of the PDM is that it is oriented to informing psychotherapy and does not address many symptoms in detail such as traumatic brain injury or neurotoxic exposure. Additionally, the PDM does not include diagnoses for traumatic brain injury or neurotoxic exposure that clinicians routinely encounter in forensic work.

The DSM is not a static document; it is continually undergoing revisions in order to keep current with clinical and research information and practices. The DSM and ICD approaches purportedly are heading in more common directions, but the other approaches to classification, such as the PDM, might offer valuable insight. Moreover, the next iterations of the diagnostic manuals need to actively consider the research available, different approaches, such as the prototypical one, and the pitfalls they represent for forensic assessment.

Conflict of Interest and Ethical Blind Spots

“There is this assumption that a tie with a company is evidence of bias. But these people [APA panel members] can be objective.”

Darrel Regier, research director for the APA, in an interview with *USA Today* in 2009.

“It is difficult to get a man to understand something when his salary depends upon his not understanding it.”

Upton Sinclair

To its credit, the APA instituted a conflict of interest policy requiring all panel members on the DSM-5 to file financial disclosure statements. Elsewhere, we reported (Cosgrove and Krinsky 2012) that this new APA requirement rendered the

DSM's disclosure policy as more congruent with most leading medical journals and federal policies on financial conflicts of interest (FCOI). DSM panel members were required to list any FCOI for 3 years prior to their appointment on the DSM, and they could not accept more than USD \$10,000 from industry (e.g., for consultancies) per year or hold more than USD \$50,000 in stock in a pharmaceutical company during their tenure on the DSM-5 task force. Unfortunately, however, transparency did not result in a reduction of FCOI of DSM-5 panel members. In fact, 69 % of DSM-5 task force members reported industry ties, which represents a 21 % increase from DSM IV task force members. Moreover, three fourths of the work groups continued to have a majority of members with ties to drug firms, and it is noteworthy that, as with the DSM IV, the most conflicted panels are those for which pharmacological treatment is the first-line intervention.

Clearly, temporarily discontinuing long-standing and substantial financial ties to industry is not a sufficient solution to prevent bias. There are ethical considerations that are relevant whether the task force (or work group) group member's involvement in the drug industry occurred prior to or after their work on the DSM. For example, knowing that one could use the prestige and power of being a task force member on the DSM-5 to leverage lucrative contracts with industry after his/her tenure on the DSM was over, could most certainly—implicitly—affect one's behavior while making decisions about revisions. Immediately after their tenure was up on the DSM-5 (i.e., after the APA approved changes), the monetary restrictions noted above no longer applied.

It should be emphasized that ethical concerns about DSM-5 panel members having commercial ties is not meant in any way to imply that any task force or work group member *intentionally* made pro-industry decisions. Decades of research have demonstrated that cognitive biases are commonplace and very difficult to eradicate, and more recent studies suggest that disclosure of financial conflicts of interest may actually *worsen* bias (Loewenstein, Sah, & Cain 2012). This is because bias is most often manifested in subtle ways unbeknownst to the researcher or clinician, and thus is usually implicit and unintentional. Physicians—like everyone else—have ethical blind spots. Social scientists have documented the fact that physicians often fail to recognize their vulnerability to commercial interests because they mistakenly believe that they are immune to marketing and industry influence (Sah and Fugh-Berman 2013). Thus, increased transparency and mandatory disclosure policies are not robust enough strategies to prevent the appearance, if not the reality, of bias in the DSM revision process. In fact, the reason the Institute of Medicine (National Research Council, 2011) recently developed new standards stating that work group members for clinical care guidelines should not have FCOI is to prevent two things—(a) the appearance of bias (and concomitantly, a decrease in public trust) and (b) the possibility of implicit, unintentional

bias. APA needs to take a stronger ethical stand and ensure that the diagnostic and clinical care guidelines comply with IOM's standards.

Ethical Guidelines and the DSM-5

We are not likely to find a diagnostic system as “unethical” *per se*, but rather find that it creates ethical concerns in its formulation and application. With the introduction of the DSM-5, we have several examples of ethical concerns.

Both psychiatrists and psychologists have ethical guidelines that emphasize the welfare of the client/patient. The Principles of Medical Ethics with Annotations Especially Applicable to Psychiatry (2013) are not as specific as the Ethical Principals of Psychologists and Code of Conduct (2010) (although for present purposes, there are sufficient similarities). Perhaps, this may be, since the study of ethics is psychological in nature and psychologists give a great deal of attention to this subject matter. For this reason, we will focus mainly on the American Psychological Association's Ethical Principals. However, the spirit of Psychiatry's Ethical Principals is very similar to that of the psychologists:

“The medical profession has long subscribed to a body of ethical statements developed primarily for the benefit of the patient. As a member of this profession, a physician must recognize responsibility to patients first and foremost, as well as to society, to other health professionals, and to self... A physician shall be dedicated to providing competent medical care with compassion and respect for human dignity and rights... A psychiatrist should not be a party to any type of policy that excludes, segregates, or demeans the dignity of any patient because of ethnic origin, race, sex, creed, age, socioeconomic status, or sexual orientation...” (The Principles of Medical Ethics With Annotations Especially Applicable to Psychiatry 2013 Edition, p.3.)

We will now review the basic Principles of the Psychologist's Ethics Code (2010) to see what dilemmas the DSM-5 may present, keeping in mind that for the most part, these ethical guidelines apply to most mental health professionals.

Standard 3. Human Relations

3.06 Conflict of Interest

“Psychologists refrain from taking on a professional role when personal, scientific, professional, legal, financial or other interests or relationships could reasonably be expected to (1) impair their objectivity, competence or effectiveness in performing their functions as psychologists.” The majority of DSM panel members were psychiatrists; this manual was developed by a specialty organization with a medical orientation. Thus, it is important to consider the ways in which guild

interests and intellectual conflicts of interest may have affected the revision of the DSM-5.

Standard 9. Assessment

9.01 Bases for Assessments

(a) “Psychologists base the opinions contained in their... diagnostic... statements... on information... sufficient to substantiate their findings. (see also Standard 2.04, Bases for Scientific and Professional Judgments).” The DSM-5 reports kappa levels as acceptable that are far below the standard level of acceptability (usually over .60; Landis & Koch, 1977). The DSM-5 does not offer sensitivity and specificity levels. Psychologists must keep this in mind when teaching or using the DSM-5 as a type of assessment guide.

Standard 2. Competence

2.04 Bases for Scientific and Professional Judgments

“Psychologists’ work is based upon established scientific and professional knowledge of the discipline.” The DSM-5 is a good guide overall, but is not a periodic table of atomic weights. It represents a particular guild’s best attempt at a classification of mental disorders. Psychologists must understand the limits of this rough guide and not treat it as a “bible of mental disorders,” to which it is commonly referred. .

Conclusions

In light of the documented concerns around “diagnostic inflation” (Bastra and Frances 2012), and the poor reliability and validity of many diagnoses, mental health professionals, particularly those providing forensic services, must exercise caution when utilizing the DSM-5. The simplistic thresholds and criteria give the illusion of a validity that does not exist. Also, insofar as diagnosis informs treatment, clinicians must be particularly careful that assigning a DSM-5 diagnosis will not lead to unnecessary pharmacological treatment and thereby expose patients to iatrogenic harm.

To conclude, for the forensic application of the DSM-5 in the area of psychological injury and law, ethics for both psychologists and psychiatrists dictate that it should be used with extreme care. It requires a better scientific foundation, beginning with its reliability and validity. This applies not only to the conditions mentioned but also to compensable psychological injuries such as pain conditions, posttraumatic stress disorder, and neurocognitive disorders, all of which are considered in this special issue on the DSM-5 and in the scientific literature (e.g., Bastra and Frances 2012). To return to a concrete example provided above, if major depression

cannot be diagnosed with reliability in the specialized settings of the field trials, and given that depression is a common psychological injury, the forensic assessor should use the DSM-5 with caution, if at all.

Prototypic descriptions, as in the ICD-10 and even more so in the PDM, provide a more ecologically based and contextual taxonomy and therefore may facilitate a more ethically sound use of diagnoses. Although the Health Insurance Portability and Accountability Act require the ICD for insurance reporting purposes, the PDM provides a more robust understanding of the whole person. Thus, the combination of ICD symptoms and a PDM description of personality structure may present the best overall taxonomy.

References

- American Psychological Association. (2013). *Ethical principles of psychologists and code of conduct*. Retrieved from <http://apa.org/ethics/code/index.aspx>
- Bastra, L., & Frances, A. (2012). Diagnostic inflation: causes and a suggested cure. *The Journal of Nervous and Mental Disease*, 200, 474–479.
- Bornstein, R. F., & Gordon, R. M. (2012). What practitioners want in a diagnostic taxonomy? Comparing the PDM with DSM and ICD. *Division/Review: A Quarterly Psychoanalytic Forum*, Fall, 6, 35.
- Brauser, D. (2012). DSM V field trials generate mixed results. *Medscape News*, May 8. Available from: <http://www.medscape.com/viewarticle/763519>. Accessed 2 April 2013.
- Cosgrove, L., & Krinsky, S. (2012). A comparison of DSM-IV and DSM V panel members’ financial associations with industry: a pernicious problem persists. *PLoS Medicine*, 9(3), e1001190. doi:10.1371/journal.pmed.1001190
- Cosgrove, L., & Wheeler, E. (2013). Drug firms, the codification of diagnostic categories, and bias in clinical guidelines. *Journal of Law and Medical Ethics*, 14(3), 644–653.
- Evans, S. C., Reed, G. M., Roberts, M. C., Esparza, P., Watts, A. D., Correia, J. M., et al. (2013). Psychologists’ perspectives on the diagnostic classification of mental disorders: results from the WHO-IUPsyS Global Survey. *International Journal of Psychology*, 48(3), 177–193. doi:10.1080/00207594.2013.804189
- Frances A. (2012). Newsflash from APA meeting: DSM 5 has flunked reliability test [Blog post]. *Psychology Today*. Available: <http://www.psychologytoday.com/blog/dsm-v-in-distress/201205/newsflash-apa-meeting-dsm-v-has-flunked-its-reliability-tests>. Accessed 30 June 2013.
- Gordon, R.M. (2007). PDM valuable in identifying high-risk patients. *The National Psychologist*, 16, 6, November/December, page 4. Available: <http://www.mmpi-info.com/pdm-valuable-identifying-high-risk-patients>
- Insel TR: Director’s Blog [Internet]. (2013). Available from: <http://www.nimh.nih.gov/about/director/2013/transforming-diagnosis.shtml>. Accessed May 2013.
- Karson, M. (2010). Bayes’ theorem and the DSM: Is it a book of definitions or a book of tests? *WebPsychEmpiricist*. Retrieved from http://wpe.info/papers_table.html. Accessed 1 June 2010
- Landis, J. R., & Koch, G. G. (1977). "The measurement of observer agreement for categorical data". *Biometrics*, 33(1), 159–174. doi:10.2307/2529310. 10.2307/2529310 . JSTOR 2529310. PMID 843571.

- Loewenstein, G., Sah, S., & Cain, D. M. (2012). The unintended consequences of conflict of interest disclosure. *JAMA*, *307*, 669–670.
- National Research Council. (2011). *Clinical practice guidelines we can trust*. Washington, DC: The National Academies Press.
- Regier, D. A., Narrow, W. E., Clarke, D. E., Kraemer, H. C., Kuramoto, S. J., Kuhl, E. A., & Kupfer, D. J. (2013). *DSM-5 field trials in the United States and Canada, Part II: test-retest reliability of selected categorical diagnoses*. *The American Journal of Psychiatry*, *170*(1), 59–70.
- Reiger quote from: M. Elias, "Conflicts of interest bedevil psychiatric drug research," *USA Today*, June 3, 2009.
- Ronson, J. (2011). Bipolar kids: victims of the 'madness industry'? *New Scientist*, *210*(2815), 44–47.
- Sah, S., & Fugh-Berman, A. (2013). Physicians under the influence: social psychology and industry marketing strategies. *Journal of Law Medicine and Ethics*, *41*(3), 665–672.
- Spitzer, R. L., Williams, J. B. W., & Endicott, J. (2012). Standards for DSM-5 reliability. *American Journal of Psychiatry*, *69*, 537.
- Task Force, P. D. M. (2006). *Psychodynamic diagnostic manual*. Silver Spring, MD: Alliance of Psychoanalytic Organizations.
- The Principles of Medical Ethics With Annotations Especially Applicable to Psychiatry 2013 Edition* Retrieved from: <http://www.psychiatry.org/practice/ethics/resources-standards>
- Westen, D., Defife, J. A., Bradley, B., & Hilsenroth, M. J. (2010). Prototype personality diagnosis in clinical practice: a viable alternative for DSM V and ICD-11. *Professional Psychology Research and Practice*, *41*(6), 482–487.
- Young, G. (2011). *Development and causality: Neo-Piagetian perspectives*. New York, NY, US: Springer Science and Business Media.