Bias in the Diagnostic and Statistical Manual 5 and Psychopathology

Akihiko Masuda, Joanne Qina'au, Michael Juberg, and Timothy Martin

Abstract

This chapter discusses bias in diagnostic decision making within the Diagnostic Statistical Manual of Mental Disorders (DSM) and its use. More specifically, we address three aspects of DSM and psychiatric diagnosis that make clinicians and researchers vulnerable to making a biased diagnostic decision. These are (a) value-laden, and yet vague, concepts of mental disorder, upon which the entire DSM nosology is based; (b) criterion biases, biases within diagnostic criteria for specific disorders themselves; and (c) clinician bias, bias unfolding in a clinician–client interaction. We then argue that great attention should be paid to the intersection of psychiatric diagnosis and diversity, where the mismatch of a client's social contingencies with a clinician's social contingencies are likely to unfold. Finally, we recommend further clarification and examination of bias in psychiatric diagnosis in order to best understand it and how diagnostic decision making may be linked to clinical competency so that we may position ourselves to better serve the clients who seek our help.

In the field of behavioral health, the meaning of bias is twofold. First, a bias involves preconception. It is one's unwarranted judgment, views, and reaction to a given individual on the basis of perceived membership in a particular social category(ies) while ignoring other category memberships and other personal attributes of that individual (Fiske, 2002; Snowden, 2003). Second, bias involves favoritism or unfairness. Bias in this way is viewed as a preconceived opinion or attitudes in favor of or against one person or group relative to another, usually in a way considered to be inequitable (Lewis-Fernández et al., 2010). Furthermore, biases, which can be held knowingly or unknowingly, are also theorized to regulate a particular action or inaction accordingly (Merino, Adams, & Hall, 2018). From a functional analytic perspective (Roche & Barnes-Holmes, 2003; Roche, Barnes-Holmes, Barnes-Holmes, & Hayes, 2001), bias can be understood functionally as a verbal antecedent or more broadly as its functional relation with

A. Masuda (✉)
Department of Psychology, University of Hawai'i at Mānoa, Honolulu, HI, USA
e-mail: amasuda4@hawaii.edu

J. Qina'au
Teachers College, Columbia University,
New York, NY, USA

M. Juberg · T. Martin
University of Hawaii at Manoa, Honolulu, HI, USA

subsequent behavior and consequences. Furthermore, whether a given judgment, belief, attitude, or reaction is unjustifiable is not determined by the private event itself: instead, it is judged socially and contextually.

The behavior of clinicians in the field of behavioral health is also subject to bias. This claim is somewhat antithetical to common views of clinician behavior as bias free and context independent (Insel et al., 2010; Poland & Caplan, 2004; Wakefield, 2007). However, as discussed extensively below, clinicians' activities, such as diagnostic decision making, conducting a psychological assessment, and working with a client in therapy, are *social behaviors* (i.e., verbal and rule-governed behaviors) that operate under particular sets of social contingencies (FitzGerald & Hurst, 2017; Hayes, Niccolls, Masuda, & Rye, 2002; Merino et al., 2018; Skinner, 1974). As such, their behaviors are also prone to bias, and studies have shown the bias-prone nature of clinician behavior.

For example, in their seminal study, Langer and Abelson (1974) demonstrated the effect of labels (i.e., "patient") on clinicians' judgments. The study compared clinicians representing two different schools of thought, behavioral and psychoanalytic, in their responses to the same video of a male interviewee. Half of each group was told that he was a "job applicant," and the other half was told that he was a "patient." Following the end of the videotape, all clinicians were asked to complete a questionnaire evaluating the interviewee. The interviewee was described as fairly well adjusted by the behavioral therapists, regardless of the label supplied. For the psychoanalytic therapists, however, when the interviewee was labeled as "patient," he was described as significantly more disturbed than when he was labeled as "job applicant." This study supports the notion that clinician behaviors are prone to bias based upon a series of preconceptions that they inadvertently draw upon and that the extent of bias may be moderated by clinician factors (e.g., the nature of behavioral health training). It has been 40 years since Langer and Abelson's groundbreaking study, and while researchers have detailed an array of bias-related errors in clinical judgment,

behavioral health professionals know relatively little about how to counterbalance the negative impact of clinician bias and bias found in psychiatric diagnosis tools (Lilienfeld & Lynn, 2015).

Present Chapter

The present chapter will address bias in psychiatric diagnosis by focusing on the Diagnostic Statistical Manual of Mental Disorders-5 (DSM-5) (American Psychiatric Association, 2013). More specially, we argue that biases are unavoidable in psychiatric diagnosis because the very behavior of diagnostic decision making is a stream of verbal and rule-governed behaviors (Hayes, Barnes-Holmes, & Roche, 2001; Skinner, 1974) that are shaped and maintained within value-laden and context-dependent social contingencies (Frances, 2013c; Lacasse, 2014; Poland & Caplan, 2004). We then argue that great attention should be paid to the intersection of psychiatric diagnosis and diversity, where the mismatch of a client's social contingencies with a clinician's social contingencies are likely to unfold (Delphin-Rittmon et al., 2015; Hunsley & Mash, 2007; Jani, Johnson, Banu, & Shah, 2016; Snowden, 2003). Broadly, we propose that understanding what DSM nosology measures and what it does not (Frances, 2013b; Lacasse, 2014), being aware of the inevitability of bias in psychiatric diagnosis, and using the DSM nosological system accordingly are important steps for promoting clinical competency in psychiatric diagnosis (Haynes & O'Brien, 2000; Haynes, Smith, & Hunsley, 2011; Poland & Caplan, 2004). Finally, we end the chapter with six actionable evidence-based recommendations to address bias in the DSM-5 and psychopathology: reposition the DSM-5 as a descriptive heuristic tool; increase the professional and public awareness of value-laden and biased nature of psychiatric diagnosis; take DSM diagnostic nosology lightly; promote cognitive flexibility, perspective taking, and empathy; clarify the goals of diagnostic assessment and potential biases in the diagnostic process; and create a safe and nonthreatening learning context.

Psychiatric Diagnosis

Psychiatric diagnosis has been a dominant topic of debate in the field of behavioral health for years. This is because psychiatric diagnosis often serves as the basis of other major clinical activities, such as case formulation, assessment, and therapy. As such, in some contexts, psychiatric diagnosis is viewed as an overarching activity of a clinician that permeates into a range of clinical activities. For example, Alarcón (2009) refers to psychiatric diagnosis as follows:

> Understood as the processing of complex information regarding symptoms, behaviors, emotional correlates and eventual neurobiological substrates by means of history-taking and actual observation of psychopathological events, psychiatric diagnosis aims at reaching a comprehensive perspective of the patient's experience, so that the most appropriate treatment can be offered, and result in clinical improvement, more efficient personal functioning, and a more comfortable quality of life of the patient and his/her family (p 131).

In other contexts, the term *psychiatric diagnosis* refers more narrowly to either the assignment of a particular diagnostic label (e.g., "Alcohol Use Disorder") to a set of symptoms or implicitly to an individual who experiences these symptoms. In the literature of biases in psychiatric diagnosis (e.g., Benson, Donnellan, & Morey, 2017; Delphin-Rittmon et al., 2015; Merino et al., 2018), psychiatric diagnosis is often referred to as the assignment of a particular diagnosis(es). Even for this latter account, the aim of the psychiatric diagnosis is clinical utility. In fact, the DSM-5 states that "The diagnosis of a mental disorder should … help clinicians to determine prognosis, treatment plan, and potential treatment outcomes for their patients" (American Psychiatric Association, 2013, p. 20).

To date, the DSM has been the standard of psychology and psychiatry in North America for the past 30 years. However, the DSM continues to pose several heated controversies (Bredström, 2017; Frances, 2013b; Gazzaniga, Heatherton, & Halpern, 2016; Lacasse, 2014). One such controversy centers around bias directed toward particular groups of individuals (Merino et al., 2018). We find this controversy to be particularly alarm-

ing as both professionals and the general public often view the DSM and the process of diagnostic decision making by a clinician to be objective and bias free (Poland & Caplan, 2004).

In response to this discrepancy, this chapter addresses three aspects of DSM and psychiatric diagnosis that make clinicians and researchers vulnerable to making a biased diagnostic decision. These are (a) *value-laden, and yet vague, concepts of mental disorder*, upon which the entire DSM nosology is based (Frances, 2013a, 2013b); (b) *criterion biases*, biases within diagnostic criteria for specific disorders themselves (Hartung & Widiger, 1998; Jane, Oltmanns, South, & Turkheimer, 2007; Widiger, 1998; Widiger & Spitzer, 1991); and (c) *clinician bias*, bias unfolding in a clinician–client interaction (e.g., Merino et al., 2018; Poland & Caplan, 2004).

We do not suggest that the field of behavioral health replace DMS-5 with an alternative, such as the Research Domain Criteria (RDoC) (Insel et al., 2010). Instead, at this point in time, we simply advocate that we use the DSM-5 wisely. More specifically, we argue that clinicians must scrutinize how the DSM defines a mental disorder and the underlying assumptions behind that definition. Additionally, we argue that clinicians take the inevitability of bias into consideration throughout the entire course of diagnostic decision making. The issue of bias is complicated because of its socially derived nature. Finally, we argue that this careful use of the DSM diagnostic system should be emphasized, especially when clinicians work with clients from sociocultural backgrounds other than their own (Masuda, 2014a, 2014b; Neighbors et al., 1999).

Ambiguity Within the Definition of Mental Disorder

As noted above, the general public continues to believe that the process of psychiatric decision making (by clinicians) and psychiatric nosology (e.g., DSM) are objective, context neutral, and bias free (Houts, 2001; Lacasse, 2014; Poland & Caplan, 2004; Wakefield, 1999). For example,

common mental disorders, such as depression (i.e., major depressive disorder) and alcoholism (i.e., alcohol use disorder), are presented to the public as *diseases with known etiologies* that are objectively identified by a biological test (Deacon, 2013; Lacasse & Leo, 2005). However, according to the DSM-5, a mental disorder is simply a set of behavioral, cognitive, and emotional symptoms (Frances, 2013b; Lacasse, 2014), and the DSM is a descriptive, not etiological, classification system (Frances, 2013a). The DSM-5 does not make reference to causality in its definition of *mental disorder* at all:

> A mental disorder is a syndrome characterized by clinically significant disturbance in an individual's cognition, emotion regulation, or behavior that reflects a dysfunction in the psychological, biological, or developmental process underlying mental functioning. Mental disorders are usually associated with significant distress or disability in social, occupational, or other important activities (American Psychiatric Association, 2013, p. 20).

It is important to note that mental disorder, being a descriptive term, is not itself the source of bias. What is problematic is that the concept of mental disorder in the DSM-5, as well as in its predecessors, is *descriptively vague* (Follette & Houts, 1996; Houts, 2001; Rogers & Mintzker, 2016). In particular, the definition of mental disorder used by the DSM-5 does not provide sufficient descriptive clarity regarding the boundaries between what is a mental disorder and what is not and how a mental disorder is different from normal distress in life (Frances, 2013c).

With the definition of mental disorder vaguely defined, clinicians are required to rely on a series of their own judgments to determine whether a client's experiences fit the criteria and threshold of a given mental disorder (Follette & Houts, 1996; Frances, 2013b; Houts, 2001; Lacasse, 2014; Lilienfeld, 2014). Take the concept of "clinically significant disturbance" as an example: according to the DSM-5, for any set of symptoms to be judged as a mental disorder, it has to co-occur with "clinically significant disturbance." For the DSM-5, it is the clinically significant disturbance that differentiates a mental disorder from normal distress in life.

However, it is challenging for clinicians to first define what disturbance is and then to judge whether disturbance associated with the set of symptoms reaches the threshold of *clinical significance*. As such, it is easy to imagine that two clinicians are likely to have different levels of diagnostic thresholds for a given disorder or for a given client. It is also easy to imagine that the diagnostic threshold of clinically significant disturbance set by a clinician drifts over time.

Harmful dysfunction (Wakefield, 1999, 2007) and maladaptiveness are constructs that are relevant to that of clinically significant disturbance. Although these terms are not included in the definition of mental disorder, they are implied in the discussion of a DSM account of mental disorder. Proponents of the DSM argue that the constructs of harmful dysfunction and maladaptiveness implied within the DSM definition of mental disorder help clinicians differentiate a mental disorder from socially deviated behaviors, as well as from normal distress in life (Ghaemi, 2013; Wakefield, 2007). However, once again, these constructs themselves do not specify any categorical parameter (e.g., diagnostic threshold or criteria) for diagnostic decision making (Follette & Houts, 1996; Houts, 2001). Furthermore, while harmful dysfunction and maladaptiveness can be identified, it is quite challenging to attribute them solely to a mental disorder while ruling out other potential precipitating or causal factors, such as social climate, economic crisis, and cultural norms.

Another area of difficulty within the construct of mental disorder is to judge whether symptoms experienced by a client are *culturally approved*. In its definition of *mental disorder*, the DSM-5 describes the role of cultural norms in diagnostic decision making as follows:

> An expectable or culturally approved response to a common stressor or loss, such as the death of a loved one, is not a mental disorder. Socially deviated behavior (e.g., political, religious, or sexual) and conflicts that are primarily between the individual and society are not mental disorder unless the deviance or conflict results from a dysfunction in the individual, as described above (American Psychiatric Association, 2013, p. 20).

According to the DSM-5, a given set of symptoms should not be judged as a mental disorder if it is judged to be culturally approved. This is the case even when the set of symptoms is associated with clinically significant disturbance. Once again, judging whether the client's symptoms are culturally approved depends heavily on the clinician's personal values and judgment, as well as his or her awareness of the client's sociocultural background and social norms (Poland & Caplan, 2004). An example of this challenge is illustrated by Merino et al. (2018) as follows:

> Consider, for example, a Black man who has grown up in a society where men and boy of colors are disproportionately targeted by law enforcement... His vigilance in everyday life might be perceived as a natural consequence of racial profiling by one provider, whereas that same behavior might be interpreted as paranoia related to schizophrenia by another… This single difference in how a provider interprets symptom presentation can dramatically alter subsequent discussions surrounding the patient's psychiatric symptoms or screening for a specific condition (p. 2).

In this exemplar, we do not necessarily argue that one of the clinicians makes a *right* or *better* diagnostic decision and the other one does not. Instead, we attempt to highlight the extreme difficulty in judging whether the behavior of a client (e.g., client's vigilance in the example above) is culturally approved or acceptable. Reading the exemplar above, many of us may feel that the first provider makes the *correct, and culturally sensitive*, diagnostic decision and the second one does not. However, depending on the circumstance and the client, the decision made by the first clinician can be extremely stereotypical and biasing (e.g., "All Black men are hypervigilant because of their upbringing in a racist environment"). Furthermore, given the nature of the DSM nosological system, as well as the reliance on the clinician's subjective judgment, we may not want to assume that there is an *ontologically correct diagnostic decision*. Diagnostic decision making is a social and interpersonal phenomenon: perhaps it is something to be justified, not to be discovered. As such, our recommendation for clinicians is to use the DSM diagnostic system wisely by explicitly acknowledging and identifying factors that contribute to their diagnostic decision making, including potential biases.

Regarding cultural considerations, it is important to acknowledge that the DSM-5 offers a guideline for the cultural formulation of psychiatric diagnosis (APA, 2013, pp. 749–759). This section is added to the DSM-5 to assist clinicians in making a correct psychiatric diagnosis and avoiding misdiagnosis. However, this effort seems to fall short because the DSM-5 does not clearly state how to use the guideline to promote an *accurate* diagnosis, for example, by taking biases into consideration in diagnostic construct, diagnostic thresholds, application of diagnostic criteria, and interpretation of assessment data (Bredström, 2017).

In sum, given this vague definition of mental disorder, clinicians need to rely on their own criteria of a mental disorder, rendering them vulnerable to making biased decisions. More specifically, a clinician may attend to certain information of the client while deemphasizing or overlooking other information somewhat subjectively (Merino et al., 2018; Whooley, 2014). By definition, this resultant process itself can be viewed as bias (Poland & Caplan, 2004). Allen Frances, the chair of the task force for the DSM-IV, and one of the most well-known psychiatrists in the world, critiques this diagnostic decision-making process for reliance on "fallible subjective judgements" (Frances, 2013c, p. 111).

Criterion Bias: Biases Within the Diagnostic Criteria Themselves

One of the controversies in the DSM-5 and the foundational predecessors from which it was developed (i.e., DSM-III-R, DSM-V, DSM-V-TR) is *criterion bias,* a bias toward particular sociocultural groups of individuals that is built into the diagnostic criteria themselves (Hartung & Widiger, 1998). This form of bias is said to occur because the DSM diagnostic criteria appear to be determined based primarily on the social

and moral standards of a particular sociocultural group.

It is important to note that taking account of group differences (e.g., racial, ethnic, or gender differences) does not in itself constitute bias. In many circumstances, responding to these differences is essential, and ignoring these differences reflects a kind of bias (Snowden, 2003). For example, although major depressive disorder (MDD) is subject to gender bias at the level of construct, not adequately acknowledging gender difference due to the types of MDD symptoms endorsed by men and women may undermine the understanding and treatment for women and men diagnosed with MDD.

It is also important to note that not all mental disorders and their diagnostic criteria are subject to criterion bias. For example, the diagnosis of enuresis, an elimination disorder in the DSM, appears to be relatively bias free, especially when behavioral criteria (i.e., repeated voiding of urine into bed or cloths at least twice a week for at least three consecutive months) and age criteria (i.e., at least five years) are met. What we will address is that, ideally, diagnostic criterion sets would be neutral for any social categorizations (e.g., gender neutral, race neutral, or ethnicity neutral). However, as the DSM stands, the criterion sets of certain mental disorders may disproportionally favor the manner in which a given disorder appears in one group more so than in another (Bredström, 2017).

Gender Criterion Biases

One such criterion bias that has been discussed extensively for the past 35 years is gender bias (Jane et al., 2007; Kaplan, 1983; Widiger & Spitzer, 1991). For example, in a discussion on the potential gender bias in the diagnosis of mental disorder in DSM-IV, Hartung and Widiger (1998) elucidate the following point:

> … an inaccurate estimate of the differential sex prevalence of a disorder will… be obtained in a fully representative epidemiologic study if the diagnostic criteria for the disorder are themselves biased in favor of one sex relative to the other (p. 267).

Gender bias in this context is a form of criterion bias that unfairly assumes that stereotypical masculine or feminine characteristics are pathological. To date, scholars have argued that personality disorders, mood disorders, conduct disorder, somatization disorder (now called somatic symptom disorder in DSM-5), schizophrenia, and schizoaffective disorder are subject to gender bias (see Hartung & Widiger, 1998; Widiger, 1998). In this section, we will review gender criterion biases in personality disorders and mood disorders.

Personality disorders that are subject to gender bias include histrionic personality disorder (HPD), borderline personality disorder (BPD), dependent personality disorder (DPD), narcissistic personality disorder (NPD), antisocial personality disorders (APD), and paranoid personality disorder (PPD). Widiger (1998) argues that gender criterion bias unfolds in diagnostic constructs, diagnostic thresholds, and applications of diagnostic criteria. Consider DPD as an example. As seen in the DSM-5, defining diagnostic features of DPD have continued to include difficulty in making everyday decisions without advice and reassurance from others and difficulty in expressing disagreement with others because of fear of loss of support or approval. As discussed extensively elsewhere (Ford & Widiger, 1989; Hartung & Widiger, 1998; Kaplan, 1983), these criterion features are known to be a stereotypically feminine form of socialization in many Western sociocultural contexts. Given the general criterion set of a mental disorder, these symptoms included in the diagnostic criteria of DPD can be viewed as culturally approved behaviors at least for some individuals in some sociocultural contexts. As such, more women diagnosed with DPD than men in Western sociocultural contexts may reflect a biased view of considering culturally regulated practice as a mental disorder (Grant et al., 2004; Trull, Jahng, Tomko, Wood, & Sher, 2010).

Regarding the application of diagnostic criteria, recent research has also shown that part of the diagnostic criteria for a disorder, not necessarily the whole diagnostic criteria, may possess gender bias (Benson et al., 2017; Hoertel et al., 2018).

For example, the chronic feeling of emptiness is a defining feature of BPD. A study shows that the experience of emptiness is related to the diagnosis of BPD, particularly for women but less so for men (Benson et al., 2017). Similarly in the diagnostic criteria of BPD, suicidal or self-mutilation behavior and affective instability are found to be relevant to women but less so to men (Hoertel, Peyre, Wall, Limosin, & Blanco, 2014). Finally, for the diagnostic criteria of NPD, lack of empathy is found to be more relevant to men than women (Hoertel et al., 2018). This *differential relevance,* not differential ratio or degree, of particular symptoms warrants careful attention. Diagnostic criteria groups should not be biased around social categorization: gender difference in this domain suggests difference in the *very construct* of these mental disorders across gender.

Another mental disorder that is subject to gender criterion bias is major depressive disorder (MDD). Epidemiological studies have consistently shown a higher prevalence of MDD in women relative to men with the ratio close to 2:1. Nevertheless, this gender imbalance in depression has been one of the major unsolved issues in psychiatric epistemology (Piccinelli & Wilkinson, 2000). Whereas there are several hypotheses for the discrepancy, criterion bias is one possible explanation. More specifically, some scholars argue that the diagnostic criteria of MDD are systematically skewed toward symptoms more frequently endorsed by women while overlooking those more frequently experienced by men (Romans, Tyas, Cohen, & Silverstone, 2007). A recent meta-analysis shows that depressed women are more likely to report depressed mood, appetite disturbance or weight change, and sleep disturbance than depressed men, while depressed men are more likely to report alcohol or drug misuses and risk taking or poor impulse control than depressed women (Cavanagh, Wilson, Kavanagh, & Caputi, 2017). Results of this meta-analysis have several key implications regarding gender biases. One such implication is the underdiagnosis of MDD in men due to fewer items within the diagnostic criteria of MDD that reflects the behavior of depressive men.

Finally, once again, it is important to clarify that taking account of group differences is not itself a bias. In many circumstances, responding to these differences is essential, and ignoring these differences reflect a kind of bias (Snowden, 2003). The DSM-5 reports that some mental disorders are more prevalent in men than women (e.g., APD and pedophilic disorder have been found to be more common in men). Gender difference in the prevalence ratio of these mental disorders, if measured accurately, is not necessarily indicative of bias at a measurement level. However, viewing these conditions as mental disorders may involve biases (e.g., subjective moral, ethical, and legal judgment) at a conceptual level.

Ethnic and Racial Criterion Biases

In addition to gender criterion bias, researchers have found the DSM-5 to have criterion bias toward particular racial and ethnic groups of individuals (Delphin-Rittmon et al., 2015; Merino et al., 2018). From a functional analytic perspective (Masuda, 2014a, 2016), ethnic- and racial-criterion bias is inevitable as the DSM criterion sets are based on the topographically defined behavioral phenomenology of particular sociocultural groups and the symptom expression of a given disorder often varies across cultures. To date, clinicians and scholars argue that ethnic- and racial-criterion bias exists most impactfully in the context of anxiety disorders, schizophrenia spectrum disorders, and personality disorders (Grant et al., 2004; Lewis-Fernández et al., 2010; Marques, Robinaugh, LeBlanc, & Hinton, 2011; Schwartz & Blankenship, 2014). In this section, due to the limited space, we will focus on ethnic- and racial-criterion bias in anxiety disorders.

In a review of culture and anxiety disorders, Lewis-Fernández and colleagues (2010) highlight "possible mismatches" between the DSM criteria and local phenomenology of particular anxiety disorders. More specifically, they identify panic disorder (PD), social phobia, social anxiety disorder (SAD), generalized anxiety disorder (GAD), agoraphobia without panic disorder (AWOPD), and obsessive-compulsive

disorder (OCD) to be subject to this mismatch. In their review, they first present cross-cultural differences in prevalence ratios of these anxiety disorders:

> Perhaps, the most striking aspect of this review is the degree of cross-cultural variability in documents in the prevalence of the anxiety disorders, even when the same diagnostic instrument is applied. Whereas US and European rates generally converge, their position relative to prevalence in other countries is not always high or low. Studies in the United States and Europe show higher prevalence rates of PD, Specific Phobia, and SAD than most other national surveys. In contrast, for AWOPD, OCD, and GAD, US and European rates fall within the international range. With some exceptions, the lowest rates are consistently found in Asia and Africa, and are usually replicated by lower rates of disorder among US populations of Asian and African descent (p. 225).

The authors then argue that the *mismatch* between the DSM criteria and the local symptom expression of these anxiety disorders may serve as the source of criterion bias:

> The cause for this degree of variability remains unclear. Although measurement limitations are likely involved, these do not necessarily invalidate concerns over lack of validity or precision in DSM-IV-TR criteria, as the two issues are intimately linked. Throughout the review, we have noted possible mismatches between the DSM criteria and the local phenomenology of the disorder in a specific cultural context (pp. 225–226).

The three most salient examples of mismatches between DSM-defined symptoms and cross-cultural conceptions of anxiety disorders presented in the review include the unexpectedness and ten-minute crescendo criteria in PD, the definition of social anxiety and social reference group in SAD, and the priority given to psychological symptoms of worry in GAD. Whether or not a PA is expected is inevitably informed by culture. As an example, if someone in Vietnam knows that he or she is prone to *trung gio´*, a locally known condition related to PAs said to be caused by the wind, then he or she may anticipate a PA on a windy day, thereby increasing his or her chances of experiencing a PA. Additionally, PAs are expected after traumatic events in many

cultures, which makes clients experiencing PAs less likely to be diagnosed with PD and more likely to be diagnosed with PTSD. The authors suggest that PD should be defined by PD symptoms, such as autonomic arousal, catastrophic cognitions, and postattack behaviors, rather than by unexpected PAs.

In the case of SAD in Japan, fear of offending others—allocentric fear—is extremely common and is also somewhat common in Western cases of SAD. Integrating this cross-cultural understanding of SAD symptomology would serve to elevate the DSM's utility across a broader range of contexts.

Similarly, in case of GAD, the authors recommend expanding criteria to include symptoms more commonly found in cultures other than Western ones, congruent with a GAD diagnosis. Somatic symptoms, such as palpitations, bowel symptoms, dizziness, and indigestion, could be added to reduce criterion biases; this recommendation is especially critical as clients who present with somatic anxiety tend to exhibit more intense levels of distress, disability, and use of medical services, relative to clients presenting with psychological symptoms (Lewis-Fernández et al., 2010). The DSM-5 does not include any of the recommendations made in the Lewis-Fernández and colleagues (2010) review. As the DSM nosology is increasingly applied to other cultural settings, there should be a careful and systematic revisiting of data that inform the current criteria.

Criterion Bias: Summary

In sum, the DSM-5 is subject to criterion bias because its diagnostic criteria seem to primarily reflect the phenomenological, social, and moral standards of a particular sociocultural group while not taking into account those of other sociocultural groups (Bredström, 2017; Pilgrim, 2014). In fact, the diagnostic criteria of mental disorders in the DSM-5 were determined by a small group of experts. Whooley (2014) describes the group decision-making processes unfolding

in the development and finalization of DSM-5 as follows:

> With the DSM we get a rare thing: experts in a room making decisions on how to define disease categories. Rarely does the social construction of disease occur so deliberately and consciously. Rarely is the logic behind certain diagnostic choices made so explicit. And rarely are the conceptual fissures so exposed (p. 94).

Clinician Bias: Functional Analytic Perspectives

Of all the biases discussed in the context of the DSM, clinician bias has been the most controversial one. As described above, a clinician's diagnostic decision making is bias laden as it relies "exclusively on fallible subjective judgements (Frances, 2013c, p. 111)." To highlight the inevitability of clinician bias in psychiatric diagnosis, Poland and Caplan (2004) describe the social contingencies that perpetuate it as follows:

> …there is a widespread assumption that therapists overcome their biases. But the assumption is a myth because biases are unavoidable … The process of trying to understand the nature of people's anguish, fears, and depression is quite complex; there are usually many unknowns, and settling on a tidy diagnosis can reduce one's anxiety. As a result, the clinician's human need to simplify the picture comes into play. This is intensified under pressures from health maintenance organizations and insurance companies or when the patient is dangerous. Thus, in making clinical diagnoses, the therapist will attend to, emphasize, and use certain information while de-emphasizing or ignoring other information, and biases shape the nature of that selectivity (p. 11).

From a functional analytic perspective, the claim that biases are unavoidable is not so surprising, given that the underlying behavioral process of biasing is ordinary human linguistic and cognitive process (Hayes et al., 2002; Masuda, Hill, Morgan, & Cohen, 2012; Roche et al., 2001). The following sections review a contemporary functional analytic model of behavioral process that underlies a bias.

Contemporary Functional Analytic Accounts of Bias, Prejudice, and Stigma

For the past 30 years, modern functional analytic accounts of complex human behaviors have been applied to social processes, such as bias, stigma, and prejudice (Hayes et al., 2002; Lillis & Levin, 2014; Masuda et al., 2012; Roche et al., 2001; Roche, Barnes-Holmes, Barnes-Holmes, Stewart, & O'Hora, 2002). One such conceptual model is relational frame theory (RFT; Hayes et al., 2001).

From an RFT perspective (Roche et al., 2001), a bias or biasing is an example of human language or cognition that has been "inappropriately" applied. Human language or cognition in this sense is a *generalized verbal operant* (i.e., verbal behavior) that operates under particular processes of a contingency. Although a detailed account of RFT is beyond the scope of the present chapter (see Hayes et al., 2001 for a full treatment), a brief summary regarding its implications for bias is warranted here.

First, a generalized verbal operant is relational. The term *relational* refers to a particular functional quality within the contingency of reinforcement. From an RFT account, a generalized verbal operant is relational in that a *response to a given event reflects the particular association of that event to other events* (Hayes et al., 2001). Let us give an example. In a social context, a generalized verbal operant allows us to respond to a given person in such a way that reflects image, attitudes, and perceptions associated with that person. When a clinician meets a given client, the clinician's perception of the encounter with that client is shaped by his or her *direct* experience with that client, as well as any preconception, stereotype, or bias that is automatically evoked by the presence of the client. If the client reminds the clinician of someone whom he or she extremely dislikes, the clinician may respond to the client in a particular way that reflects that previous personal experience of dislike (Poland & Caplan, 2004). If a clinician attends to a certain demographic feature of a client (e.g., ethnic background and neighborhood) more so than other information, the clinician's personal history

relevant to that demographic information may shape the way the clinician interprets the client's statement during a diagnostic assessment. This relational feature of generalized verbal operant (i.e., responding to one in terms of its relation to another) reflects the defining feature of bias.

Second, the relational features of the generalized verbal operant are *automatically derived*, oftentimes without awareness. This derived quality is an essential feature of human language and cognition that differentiates it from other forms of operant behavior (Hayes et al., 2001). In psychological science literature, this derived learning is often referred to as latent learning, indirect learning, or learning without direct reinforcement (Roche et al., 2001). This derived quality of verbal behavior is also acknowledged in the literature of attitude formation and implicit vs. explicit attitudes or biases in social psychology.

It is important to note that given the involuntary or automatically derived nature of implicit attitudes, one's implicit attitudes often contradict with explicit attitudes. For example, while explicitly opposing or denying any prejudiced attitudes, individuals often demonstrate implicit racial bias toward a particular ethnic or racial group of individuals (Green et al., 2007; Greenwald, McGhee, & Schwartz, 1998). These implicit attitudes seem to reflect well-established and general schemas (i.e., relational network) that have been steadily shaped throughout one's history and maintained independently from explicitly stated attitudes. Extant literature shows that these derived implicit biases also regulate clinical decision making in psychiatric diagnostic procedures for a particular client (Merino et al., 2018).

Third, a bias is often unavoidable as it can occur in virtually every context. This is because the occurrence of a generalized verbal operant is not restricted to the physical properties of the environmental context (Hayes et al., 2001, 2002). For example, the occurrence of taking an illicit drug is limited to the environment where that illicit drug is available. On the other hand, the behavior of biasing can occur in every waking moment regardless of the situation.

Fourth, RFT suggests that various forms of bias (e.g., racial bias, gender bias, sexuality bias), although varying in content, may not be qualitatively distinct from one another in process (Levin et al., 2016; Lillis & Levin, 2014). This conceptual position is supported in part by the finding that prejudiced biases and attitudes toward various groups tend to co-occur and comprise a single latent variable (Bäckström & Björklund, 2007). Evidence shows that individuals who are prejudiced toward Black Americans are also likely to be biased against other groups, such as other ethnic minorities, women, and sexual minorities (Akrami, Ekehammar, & Bergh, 2011). The applied implication of this claim is that we should target this meta- or generalized process, as opposed to each distinct form of bias, when increasing awareness of bias in the context of psychiatric decision making.

Fifth, according to RFT, bias, stigma, and prejudice are also inherently rigid once they are formed (Hayes et al., 2002). New ideas are met with resistance when they are not aligned with stereotype-consistent beliefs (Moxon, Keenan, & Hine, 1993; Watt, Keenan, Barnes, & Cairns, 1991), and efforts to change unwanted thoughts often paradoxically increase their frequency and intensity (Wegner, 1994; Wenzlaff & Wegner, 2000). Even if the expression of a biased attitude is extinguished in a particular context, it still remains in a person's repertoire (Wilson, Lindsey, & Schooler, 2000).

Finally, as implied throughout the chapter, the behavior of biasing is socially shaped by the member of the sociocultural community (Hayes & Brownstein, 1986; Skinner, 1957, 1974). Evidence shows that the behavior of biasing is developed and reinforced early in childhood and continues throughout one's lifetime (Baron & Banaji, 2006; Hayes et al., 2001; Pauker, Ambady, & Apfelbaum, 2010; Pauker, Williams, & Steele, 2016). This may be because across many cultural contexts, biases allow humans to navigate themselves more easily through complex sociocultural interactions (Macrae, Bodenhausen, Milne, & Jetten, 1994).

Clinician Biases: Exemplars

The previous section addresses the underlying behavioral processes of clinician bias. In this section, we are going to present specific examples of clinician biases manifested in the context of psychiatric diagnosis.

In their comprehensive review of clinician bias, Poland and Caplan (2004) identify several exemplars of how ordinary human cognitive and behavioral processes unfold as biases in the context of psychiatric diagnosis. One such exemplar is the clinician's general attitudes and beliefs about certain groups of individuals. These attitudes and beliefs may reflect some of the explicit and implicit stereotypical notions of particular groups, such as "Asian Americans are a model minority" (Chou & Feagin, 2015) and "Black men don't like to work" (Way & Rogers, 2015). These implicit and explicit attitudes toward particular groups of individuals may influence a clinician's diagnostic decision for a given client even when a standardized diagnostic tool, such as the Structured Clinical Interview for DSM, is used.

In the context of behavioral health in the United States, research has continued to report that particular groups of individuals are disproportionately diagnosed with certain mental disorders (Delphin-Rittmon et al., 2015; Schwartz & Blankenship, 2014). One such mental disorder is schizophrenia. Research continuously shows that Black Americans are three to four times more likely to be diagnosed with psychotic disorders compared to White Americans (Schwartz & Blankenship, 2014). There are several explanations for these differential rates, such as the collection of situational stressors (e.g., safety, employment status) that are differentially more salient to Black American clients. In addition to those factors, clinician bias is also said to contribute to this differential ratio (Merino et al., 2018). More specifically, advocates for clinician bias suggest that presenting symptoms are interpreted differently by clinicians, depending on the racial and ethnic background of the client. For example, some authors suggest that, in part, due to the cli-

nicians' implicit and explicit preconceptions toward Black Americans, socially deviant and disruptive behavior tends to be interpreted as psychotic symptoms when presented by members of this group (Schwartz & Blankenship, 2014).

Similarly, clinician biases toward particular ethnic groups are also reported in the diagnosis of mood disorders (Delphin-Rittmon et al., 2015) and personality disorders (Grant et al., 2004). A cautionary note here is that the majority of clinicians in these studies were White Americans. As such, it is unclear whether clinician biases against Black Americans are observed among ethnic minority clinicians or whether, conversely, ethnic minority clinicians may have biased views against White American clients.

On a related note, Poland and Caplan (2004) also suggest that clinician bias is especially likely when the clinician's cultural background differs from those of clients, and this is the second exemplar. From a functional analytic perspective, the collection of information and the development of clinical understanding essentially depend on the features of the clinician–client interaction. Cultural differences in this sense are indicative of differences not in racial and ethnic characteristics but in social norms and sociocultural contingencies that maintain these norms. Interpretation of a behavioral sign (e.g., "My ancestors always guide me") using one social norm may be quite different from one made using another cultural norm (Merino et al., 2018; Sue, Fujino, Hu, Takeuchi, & Zane, 1991; Sue & Zane, 1987).

Third, according to Poland and Caplan (2004), well-known cognitive tendencies are also pervasive in the field of behavioral health, and they are often manifested as the form of clinician biases in psychiatric diagnosis. These general cognitive tendencies are as follows:

- confirmation bias,
- availability bias,
- stereotype-based memory bias,
- illusory correlations,
- halo effects,
- anchoring effects,
- actor-observer bias.

Confirmation bias is our strong tendency to place greater importance on evidence that supports our own existing beliefs while downplaying evidence that does not match what we believe. For example, when a clinician believes that men are less likely to be diagnosed with a major depressive disorder (MDD) than women, he or she may be less likely to interpret sadness or lack of interest experienced by a male client as a symptom of MDD.

Availability bias is our propensity to give priority to information that is highly salient or more easily remembered. This form of bias is problematic as it may give rise to a mental shortcut of diagnostic decision making (e.g., MDD) by merely attending to one form of information (e.g., "the client appears to be extremely sad during the initial interview") while neglecting to assess other relevant information.

Stereotype-based memory bias is the inclination to recall information that is stereotype confirming to a given client even when the information is not applied to that client. Similarly, *illusory correlation* is the propensity to see a significant association between an observed characteristic and an unobserved characteristic because that association is prevalent in some contexts.

Halo effect is the tendency to assign positive or negative traits to a client who exhibits other desirable or undesirable traits, even though the traits are not correlated. Stereotype-based memory bias, illusory correlation, and halo effect are good examples of the derived and relational nature of human language and cognition. When a clinician attends to one form of information of a client, its associated information automatically and unconsciously become available to the clinician, although the derived information may not be applicable to the client.

In the context of psychiatric diagnosis, *anchoring effect* is a clinician's tendency to outweigh the first impression of a given client over subsequently collected information in clinical decision making. Finally, the *actor-observer bias* is the tendency to see one's own problem or the problems of someone with whom one identifies as resulting from situational factors while regarding the problems of others, especially those with whom one does not identify, as resulting from an intrinsic cause. The actor-observer bias may be particularly relevant to clinician bias as this general tendency is said to occur when a clinician's sociocultural background is significantly different from that of a client in a particular way (Poland & Caplan, 2004).

Clinician Bias: Summary

In sum, this section argues that ordinary human linguistic and cognitive process of clinicians can manifest as clinician bias in the context of psychiatric diagnosis. Human linguistic and cognitive processes are prone to obscuring a clinician's direct experience with a given client by automatically evoking stereotypical attitudes and images associated with the client, oftentimes without his or her awareness. Furthermore, the sociocultural contexts of clinicians, like those of the general public, maintain such linguistic and cognitive practices, or at least do not extinguish them.

Summary and Recommendations

The present chapter addresses three primary biases within the context of psychiatric diagnosis: *value-laden, and yet vague, concepts of mental disorder*; *criterion biases*; and *clinician bias*. We argue that diagnostic decision making is highly vulnerable to bias because clinicians must rely on their fallible subjective judgments without a clear and reliable diagnostic guideline. From a functional analytic perspective, bias within this context (i.e., variability in decision making) is unavoidable as the act of diagnostic decision making is a generalized verbal operant regulated under ambiguous verbal control (Barnes-Holmes, Hayes, Barnes-Holmes, & Roche, 2001; Roche et al., 2001; Skinner, 1974).

As a solution for bias in psychiatric diagnosis, many scholars and clinicians appear to implicitly advocate making an *accurate diagnosis* by minimizing biases. Whereas this clinical recommendation is sound and well intentioned, we must carefully examine its underlying assumptions, its intended outcomes, and the feasibility of achieving this goal. At least four assumptions are made when we advocate making accurate diagnostic decisions by minimizing biases.

The first assumption is of an ontological nature: we are presuming that there is an *absolutely correct* diagnostic picture for a given clinical case, and the clinician's job is to discover it. Relevant to the topic of the present chapter, this assumption also implies that although there is a true diagnostic picture, biases and other factors prevent it from being discovered. This may be the case for some mental disorders, such as enuresis and encopresis, but it is difficult to maintain this assumption for the diagnosis of most other mental disorders, such as anxiety and mood disorders, which require subjective judgments.

The second assumption is that we can objectively identify a bias and clearly differentiate it from unbiased attitudes and beliefs. Throughout the present chapter, we have argued that psychiatric decision making is subjective and value laden. We argue that the same dilemma is also applied to the case of bias. Judging whether a given attitude or cognitive process is a bias is extremely subjective, and it depends on one's subjective point of view. Just as there is no objective test to identify a mental disorder, there is no objective method for bias. It is important to note that our intention here is not to downplay the significance of bias; what we usually refer to as biases certainly contribute to behavioral health services disparities. Rather, we attempt to elucidate the complex and nuanced nature of bias as it relates to decision making.

The third assumption, which is particularly aligned with the recommended efforts to minimize bias to make an accurate diagnosis, regards the ontological status of mental disorders. When biases within psychiatric diagnosis are discussed, experts often assume that a mental disorder is a concrete entity (e.g., medical disease) that exists in somewhat consistent form across all areas of the world while in reality, its symptom presentation may vary substantially cross-culturally. In this ontological assumption, a bias is viewed as an error or variability that obscures accurate diagnostic decision making (e.g., "misdiagnosis = bias + accurate diagnosis").

Finally, the pursuit of an accurate—absolutely correct—diagnosis is based on the assumption that doing so is absolutely necessary for greater clinical competency for determining prognosis, a treatment plan, and evaluating treatment effectiveness. We partially agree with this assumption in that psychiatric diagnosis has been part of our routine practice from the very nascent phases of the field and the DSM nosological system provides a common language across various stakeholders in the context of behavioral healthcare. However, we also argue that its clinical utility for developing and conducting effective treatment is questionable (see Follette & Houts, 1996). From a functional analytic perspective, the DSM diagnostic system is incomplete for treatment because it does not inform clinicians of any functional relations between the client's presenting concerns and their maintaining factors. Such functional understandings are crucial for formulating case conceptualization, treatment plan, and treatment evaluation and modification (Hayes & Hofmann, 2018).

The truth is that the DSM-5 is subject to many controversies, and the descriptive and value-laden nature of the DSM-5's definition of a mental disorder and subsequent diagnostic criteria are merely a few examples of such disputes. Jeffrey Lacasse (2014) uses the term *conjecture* to identify nine notable issues addressed by the critics of DSM (e.g., Frances, 2013b; Pilgrim, 2014):

- Conjecture 1: The DSM-5 definition of mental disorder is inadequate.
- Conjecture 2: DSM-5's claim that all mental disorders are medical diseases is unsupported.
- Conjecture 3: The DSM-5 is more political and less transparent than previous editions.
- Conjecture 4: The DSM-5 is unreliable.
- Conjecture 5: The ramifications of unreliable diagnosis are significant.
- Conjecture 6: The accuracy of knowledge dissemination regarding psychiatric diagnosis is poor.
- Conjecture 7: The primary utility of the DSM continues to be financial, not scientific.
- Conjecture 8: Applying DSM-5 diagnoses to clients can cause harm.
- Conjecture 9: There are viable alternatives to conventional diagnosis.

With vague value-laden concepts of mental disorders to work from, a host of criterion biases plaguing our diagnostic tools, and an array of clinician biases to contend with, how might we utilize what we know to provide fair and useful diagnoses to clients?

Recommendations

When considering how best to address the issue of bias in the DSM, the overarching recommendation is to use the DSM-5 diagnostic system wisely. As its clinical utility has come under critical scrutiny since its inception, the DSM-5 requires a more nuanced and critical understanding of its strengths and weaknesses. As such, we offer the following six recommendations.

Reposition the DSM-5 as a descriptive heuristic tool It is of vital importance for behavioral health stakeholders and the general public to know how the DSM authors came to define modern mental disorders and the diagnostic criteria of a specific mental disorder (Frances, 2013b; Lacasse, 2014). As such, our first specific recommendation to counteract bias in DSM

diagnosis is to use DSM-5 diagnostic categories *descriptively* rather than etiologically or ontologically.

Increase the awareness of value-laden and the biased nature of psychiatric diagnosis Our second recommendation is to increase the awareness among behavioral health stakeholders and the general public that psychiatric diagnosis relies on potentially biased tools and fallible subjective judgment.

Awareness of personal biases is a vital step toward mitigating their deeply ingrained patterns. To this point, enhancing awareness of one's own cognitive processes, both implicit and explicit, is a cornerstone of multicultural competency training in the field of counseling psychology (Sue, Zane, Hall, & Berger, 2009). Also, raising awareness about the gap between implicit and explicit bias (Monteith & Mark, 2005) has great potential as a first step in evidence-informed strategies to address cognitive bias. For example, encouraging awareness of one's own implicit and explicit cognitive process (e.g., prejudicial attitudes) can reduce discriminatory behavior among individuals with low-explicit and high-implicit prejudice (Son Hing, Li, & Zanna, 2002).

Although the field of behavioral health in the United States is still very much in the era of the DSM, where the medical industry markets mental disorders as medical diseases, the important issue of bias can be explored, and potentially assuaged, through a variety of mediums. One such arena is in graduate training programs (Poland & Caplan, 2004). When future behavioral health professionals enter into a graduate training program (e.g., Ph.D. in Clinical Psychology), they may already assume that a mental disorder is a medical disease, and yet few graduate training programs address the subjective nature of clinical decision making in psychiatric diagnosis (Poland & Caplan, 2004). In their six-point framework for integrating implicit bias training into training for health professionals, Sukhera and Watling (2018) suggest the following:

1. creating a safe and nonthreatening learning context,
2. increasing knowledge about the science of implicit bias,
3. emphasizing how implicit bias influences behaviors and patient outcomes,
4. increasing self-awareness of existing implicit biases,
5. improving conscious efforts to undermine the behavior regulatory impact of implicit bias,
6. enhancing awareness of how implicit bias influences others.

This training framework, which is designed to be generic to implicit biases in the field of behavioral health, can be easily tailored to the issues associated with bias in psychiatric diagnosis. Although a detailed account of this training framework is beyond the scope of this chapter, it is worthwhile to note that its guiding conceptual framework and recommended training targets, such as metacognition and mindfulness, are consistent with implications and recommendations suggested by a functional analytic account of social categorization and prejudice (Masuda, Donati, Schaefer, & Hill, 2015; Masuda et al., 2012).

Take DSM diagnostic nosology lightly Our third recommendation is to take a DSM psychiatric diagnosis lightly in clinical decision-making. Many controversies concerning the DSM-5 stem from placing a greater weight on DSM-5 diagnosis in clinical decision making than we should. When we take the DSM lightly, such as using a psychiatric diagnosis as an entry point for more thorough assessment and treatment plan, the negative impact of bias unfolding in the context of DSM psychiatric diagnosis also decreases.

Promote cognitive flexibility, perspective taking, and empathy From a functional analytic perspective (Masuda et al., 2012), the problem of cognitive bias is not in its content but its discriminative function that limits response options (leading to reach a particular decision, while alternative decisions are also plausible). Cognitive flexibil- ity, perspective taking, and empathy in the context of psychiatric diagnosis are behavioral repertoires of (a) becoming aware of one's decision-making process while examining alternative hypotheses, (b) viewing the client's psychiatric experience from the standpoint of the client, and (c) becoming connected to the sense of what it is like to be that client for seeking contextual information that may be relevant to the client's presenting concern.

Clarify the goals of diagnostic assessment and potential biases in the diagnostic process From a functional analytic perspective, whether a certain diagnostic decision is biased or not is contextually and socially determined. As such, in order to avoid further chaos caused by relativism, it is useful to clearly state the purpose of psychiatric diagnosis in a given clinical case, the process of reasoning in psychiatric diagnosis, and potential implicit and explicit cognitive biases (Roche & Barnes-Holmes, 2003; Roche et al., 2001; Roche et al., 2002). The diagnostic information can be used for various purposes, ranging from simply labeling and categorizing a person to implying the severity of current psychological condition and predicting future behavior in a legal context. Given the social nature of psychiatric diagnosis, psychiatric decision making is something to be justified, not discovered. The clearly stated goal of the psychiatric diagnosis behavioral health stakeholders and the general public to examine the extent to which the psychiatric decision and process of reasoning made by a clinician is justified.

It is also important for clinicians to document the process of reasoning in psychiatric decision making, including potential biases in a given clinical case. For example, it may be beneficial to address the extent of cultural match between the client and the clinician, as well as common cognitive biases (e.g., anchoring effect) that might have influenced a diagnostic decision in a diagnostic report. While clinician bias has been a central topic in the field of behavioral health, evidence remains limited regarding when and

how bias unfolds in the course of psychiatric diagnosis and other clinical activities (Merino et al., 2018). It is therefore extremely important to continue to examine—and document—how we best conceptualize and target bias in the context of psychiatric diagnosis (Lilienfeld, 2017).

Create a safe and nonthreatening learning context Finally, bias reduction training is often tricky to successfully perform (Twohig, Domenech Rodriguez, & Enno, 2014). It often evokes unwanted effects and attitudes in both training facilitators and participants. When teaching about bias, stereotyping, and prejudice, there should be explicit recognition that the removal of all biases is impossible, although one can change the ways to relate to them (see Sukhera & Watling, 2018, for a mindful and metacognitive awarenss of bias). Collectively acknowledging the inevitability of bias in a training group allows the group members to learn alternative ways to relate to their biases without being defensive or reactive.

In a bias reduction training, it is also important to be mindful of one's own sense of self-righteousness (Masuda, 2014a). Once again, it is the sense of self-righteousness that makes us become dismissive, defensive, and reactive toward others. Finally, it is also important to acknowledge that "remedies" can promote biases further. For example, taking a cultural diversity class may promote a stereotypical belief toward a given person (e.g., "Asians are collectivists, and Aki is an Asian, therefore, Aki must be collectivist"). To minimize the effect of bias on our clinical decision making, it is important to become aware of these very cognitive process when we engage in that activity.

Conclusion

The present chapter has discussed bias in psychiatric diagnosis by focusing on the Diagnostic Statistical Manual of Mental Disorders (DSM) (American Psychiatric Association, 2013). More specially, we argue that biases in psychiatric

diagnosis are unavoidable because the very behavior of diagnostic decision making is a stream of verbal and rule-governed behaviors that are socially shaped and maintained under a vague set of rules (Frances, 2013c; Lacasse, 2014; Poland & Caplan, 2004). We then argue that great attention should be paid to the intersection of psychiatric diagnosis and diversity, where the mismatch of a client's social contingencies with a clinician's social contingencies is likely to unfold (Delphin-Rittmon et al., 2015; Hunsley & Mash, 2007; Jani et al., 2016; Snowden, 2003). Furthermore, we propose that we take diagnostic decisions drawn from the DSM nosology with vigilance as such conclusions are inevitably value laden and bias prone (Frances, 2013b; Lacasse, 2014). Finally, we recommend further clarification and examination of bias in psychiatric diagnosis in order to best understand it and how it may be linked to clinical competency so that we may position ourselves to better serve the clients and patients who seek our help.

References

Akrami, N., Ekehammar, B., & Bergh, R. (2011). Generalized prejudice:Common and specific components. *Psychological Science, 22*(1), 57–59. https://doi.org/10.1177/0956797610390384

Alarcón, R. D. (2009). Culture, cultural factors and psychiatric diagnosis: Review and projections. *World Psychiatry, 8*(3), 131–139.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub, Washington, DC.

Bäckström, M., & Björklund, F. (2007). Structural modeling of generalized prejudice: The role of social dominance, authoritarianism, and empathy. *Journal of Individual Differences, 28*(1), 10–17.

Barnes-Holmes, Y., Hayes, S. C., Barnes-Holmes, D., & Roche, B. (2001). Relational frame theory: A post-Skinnerian account of human language and cognition. *Advances in Child Development and Behavior, 28*, 101–138. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11605362

Baron, A. S., & Banaji, M. R. (2006). The development of implicit attitudes: Evidence of race evaluations from ages 6 and 10 and adulthood. *Psychological Science, 17*(1), 53–58.

Benson, K. T., Donnellan, M. B., & Morey, L. C. (2017). Gender-related differential item functioning in DSM-IV/DSM-5-III (alternative model) diagnostic

criteria for borderline personality disorder. *Personality Disorders: Theory, Research, and Treatment, 8*(1), 87–93.

Bredström, A. (2017). Culture and context in mental health diagnosing: Scrutinizing the DSM-5 revision. *Journal of Medical Humanities.* https://doi.org/10.1007/s10912-017-9501-1

Cavanagh, A., Wilson, C. J., Kavanagh, D. J., & Caputi, P. (2017). Differences in the expression of symptoms in men versus women with depression: A systematic review and meta-analysis. *Harvard Review of Psychiatry, 25*(1), 29–38.

Chou, R. S., & Feagin, J. R. (2015). *Myth of the model minority: Asian Americans facing racism*. Routledge, New York.

Deacon, B. J. (2013). The biomedical model of mental disorder: A critical analysis of its validity, utility, and effects on psychotherapy research. *Clinical Psychology Review, 33*(7), 846–861.

Delphin-Rittmon, M. E., Flanagan, E. H., Andres-Hyman, R., Ortiz, J., Amer, M. M., & Davidson, L. (2015). Racial-ethnic differences in access, diagnosis, and outcomes in public-sector inpatient mental health treatment. *Psychological Services, 12*(2), 158.

Fiske, S. T. (2002). What we know now about bias and intergroup conflict, the problem of the century. *Current Directions in Psychological Science, 11*(4), 123–128.

FitzGerald, C., & Hurst, S. (2017). Implicit bias in healthcare professionals: A systematic review. *BMC Medical Ethics, 18*(1), 19.

Follette, W. C., & Houts, A. C. (1996). Models of scientific progress and the role of theory in taxonomy development: A case study of the DSM. *Journal of Consulting and Clinical Psychology, 64*(6), 1120–1132. https://doi.org/10.1037/0022-006X.64.6.1120

Ford, M. R., & Widiger, T. A. (1989). Sex bias in the diagnosis of histrionic and antisocial personality disorders. *Journal of Consulting and Clinical Psychology, 57*(2), 301.

Frances, A. (2013a). The new crisis of confidence in psychiatric diagnosis. *Annals of Internal Medicine, 159*(3), 221–222.

Frances, A. (2013b). The past, present and future of psychiatric diagnosis. *World Psychiatry, 12*(2), 111–112. https://doi.org/10.1002/wps.20027

Frances, A. (2013c). *Saving normal: An insider's revolt against out-of-control psychiatric diagnosis, DSM-5, Big Pharma, and the medicalization of ordinary life*. New York, NY: William Morrow &.

Gazzaniga, M., Heatherton, T., & Halpern, D. (2016). *Psychological science* (5th ed.). WW Norton & Company, New York.

Ghaemi, S. N. (2013). Requiem for DSM. *Psychiatric Times, 30*(7), 16.

Grant, B. F., Hasin, D. S., Stinson, F. S., Dawson, D. A., Chou, S. P., Ruan, W., & Pickering, R. P. (2004). Prevalence, correlates, and disability of personality disorders in the United States: Results from the national epidemiologic survey on alcohol and related

conditions. *The Journal of Clinical Psychiatry, 65*, 948.

Green, A. R., Carney, D. R., Pallin, D. J., Ngo, L. H., Raymond, K. L., Iezzoni, L. I., & Banaji, M. R. (2007). Implicit bias among physicians and its prediction of thrombolysis decisions for black and white patients. *Journal of General Internal Medicine, 22*(9), 1231–1238.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74*(6), 1464–1480.

Hartung, C. M., & Widiger, T. A. (1998). Gender differences in the diagnosis of mental disorders: Conclusions and controversies of the DSM-IV. *Psychological Bulletin, 123*(3), 260–278. Retrieved from http://www.sciencedirect.com/science/article/pii/S147101531200089X

Hayes, S. C., Barnes-Holmes, D., & Roche, B. (2001). *Relational frame theory: A post-Skinnerian account of human language and cognition*. New York, NY: Kluwer Academic/Plenum Publishers.

Hayes, S. C., & Brownstein, A. J. (1986). Mentalism, behavior-behavior relations, and a behavior-analytic view of the purposes of science. *The Behavior Analyst, 9*(2), 175–190. Retrieved from http://ezproxy.gsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=1988-31398-001&site=ehost-live

Hayes, S. C., & Hofmann, S. G. (Eds.). (2018). *Process-based CBT: The science and core clinical competencies of cognitive behavioral therapy*. Oakland, CA: New Harbinger Publications.

Hayes, S. C., Niccolls, R., Masuda, A., & Rye, A. K. (2002). Prejudice, terrorism and behavior therapy. *Cognitive and Behavioral Practice, 9*(4), 296–301. https://doi.org/10.1016/s1077-7229(02)80023-2

Haynes, S. N., & O'Brien, W. H. (2000). *Principle and practice of behavioral assessment*. New York, NY: Kluwer Academic/Plenum Publication.

Haynes, S. N., Smith, G. T., & Hunsley, J. D. (2011). *Scientific foundations of clinical assessment*. Taylor & Francis, New York.

Hoertel, N., Peyre, H., Lavaud, P., Blanco, C., Guerin-Langlois, C., René, M., … Limosin, F. (2018). Examining sex differences in DSM-IV-TR narcissistic personality disorder symptom expression using Item Response Theory (IRT). *Psychiatry Research, 260*, 500–507. https://doi.org/10.1016/j.psychres.2017.12.031

Hoertel, N., Peyre, H., Wall, M. M., Limosin, F., & Blanco, C. (2014). Examining sex differences in DSM-IV borderline personality disorder symptom expression using Item Response Theory (IRT). *Journal of Psychiatric Research, 59*, 213–219. https://doi.org/10.1016/j.jpsychires.2014.08.019

Houts, A. C. (2001). Harmful dysfunction and the search for value neutrality in the definition of mental disorder:

Response to Wakefield, part 2. *Behaviour Research and Therapy, 39*(9), 1099–1132.

Hunsley, J., & Mash, E. J. (2007). Evidence-based assessment. *Annual Review of Clinical Psychology, 3*(1), 29–51. https://doi.org/10.1146/annurev.clinpsy.3.022806.091419

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., … Wang, P. (2010). Research domain criteria (RDoC): Toward a new classification framework for research on mental disorders. *American Journal of Psychiatry, 167*(7), 748–751.

Jane, J. S., Oltmanns, T. F., South, S. C., & Turkheimer, E. (2007). Gender bias in diagnostic criteria for personality disorders: An item response theory analysis. *Journal of Abnormal Psychology, 116*(1), 166.

Jani, S., Johnson, R. S., Banu, S., & Shah, A. (2016). Cross-cultural bias in the diagnosis of borderline personality disorder. *Bulletin of the Menninger Clinic, 80*(2), 146–165.

Kaplan, M. (1983). A woman's view of DSM-III. *American Psychologist, 38*(7), 786–792.

Lacasse, J. R. (2014). After DSM-5: A critical mental health research agenda for the 21st century Sage Publications Sage CA: Los Angeles, CA.

Lacasse, J. R., & Leo, J. (2005). Serotonin and depression: A disconnect between the advertisements and the scientific literature. *PLoS Medicine, 2*(12), e392.

Langer, E. J., & Abelson, R. P. (1974). A patient by any other name...: Clinician group difference in labeling bias. *Journal of Consulting and Clinical Psychology, 42*(1), 4.

Levin, M. E., Luoma, J. B., Vilardaga, R., Lillis, J., Nobles, R., & Hayes, S. C. (2016). Examining the role of psychological inflexibility, perspective taking, and empathic concern in generalized prejudice. *Journal of Applied Social Psychology, 46*(3), 180–191.

Lewis-Fernández, R., Hinton, D. E., Laria, A. J., Patterson, E. H., Hofmann, S. G., Craske, M. G., … Liao, B. (2010). Culture and the anxiety disorders: Recommendations for DSM-V. *Depression and Anxiety, 27*(2), 212–229.

Lilienfeld, S. O. (2014). The Research Domain Criteria (RDoC): An analysis of methodological and conceptual challenges. *Behaviour Research and Therapy, 62*, 129–139.

Lilienfeld, S. O. (2017). Microaggressions: Strong claims, inadequate evidence. *Perspectives on Psychological Science, 12*(1), 138–169. https://doi.org/10.1177/1745691616659391

Lilienfeld, S. O., & Lynn, S. J. (2015). Errors/biases in clinical decision making. In R. L. Cautin & S. O. Lilienfeld (Eds.), *The encyclopedia of clinical psychology* (Vol. 2, pp. 1098–1107). Hobooken, NJ: Wiley.

Lillis, J., & Levin, M. (2014). Acceptance and mindfulness for undermining prejudice. In A. Masuda (Ed.), *Mindfulness and acceptance in multicultural competency: A contextual approach to sociocultural diversity in theory and practice. Oakland, CA: New Harbinger* (pp. 181–196). Oakland, CA: Context Press.

Macrae, N. C., Bodenhausen, G. V., Milne, A. B., & Jetten, J. (1994). Out of mind but back in sight: Stereotypes on the rebound. *Journal of Personality and Social Psychology, 67*(5), 808–817.

Marques, L., Robinaugh, D. J., LeBlanc, N. J., & Hinton, D. (2011). Cross-cultural variations in the prevalence and presentation of anxiety disorders. *Expert Review of Neurotherapeutics, 11*(2), 313–322. https://doi.org/10.1586/ern.10.122

Masuda, A. (2014a). Psychotherapy in cultural context. In A. Masuda (Ed.), *Mindfulness and acceptance in multicultural competency: A contextual approach to sociocultural diversity in theory and practice* (pp. 39–55). Oakland, CA: New Harbinger Publication.

Masuda, A. (Ed.). (2014b). *Mindfulness and acceptance in multicultural competency: A contextual approach to sociocultural diversity in theory and practice*. Oakland, CA: Context Press/New Harbinger Publications.

Masuda, A. (2016). Principle-based cultural adaptation of cognitive behavior therapies: A functional and contextual perspective as an example. *Japanese Journal of Behavior Therapy, 42*(1), 11–19. Retrieved from http://ezproxy.gsu.edu:2048/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyhref&AN=JJBT.CD.ACG.MASUDA.ACTAJC&site=ehost-live

Masuda, A., Donati, M. R., Schaefer, L. W., & Hill, M. L. (2015). Terrorism as an act in context. In M. Taylor, J. Roach, & K. Pease (Eds.), *Evolutionary psychology and terrorism* (pp. 124–148).

Masuda, A., Hill, M. L., Morgan, J., & Cohen, L. L. (2012). A psychological flexibility-based intervention for modulating the impact of stigma and prejudice: A descriptive review of empirical evidence. *Psychology, Society & Education, 4*(2), 211–223. Retrieved from.

Merino, Y., Adams, L., & Hall, W. J. (2018). Implicit bias and mental health professionals: Priorities and directions for research. *Psychiatric Services, 0*(0), appi.ps.201700294. https://doi.org/10.1176/appi.ps.201700294

Monteith, M. J., & Mark, A. Y. (2005). Changing one's prejudiced ways: Awareness, affect, and self-regulation. *European Review of Social Psychology, 16*(1), 113–154.

Moxon, P. D., Keenan, M., & Hine, L. (1993). Gender-role stereotyping and stimulus equivalence. *The Psychological Record, 43*(3), 381–394.

Neighbors, H. W., Trierweiler, S. J., Munday, C., Thompson, E. E., Jackson, J. S., Binion, V. J., & Gomez, J. (1999). Psychiatric diagnosis of African Americans: Diagnostic divergence in clinician-structured and semistructured interviewing conditions. *Journal of the National Medical Association, 91*(11), 601–612.

Pauker, K., Ambady, N., & Apfelbaum, E. P. (2010). Race salience and essentialist thinking in racial stereotype development. *Child Development, 81*(6), 1799–1813. https://doi.org/10.1111/j.1467-8624.2010.01511.x

Pauker, K., Williams, A., & Steele, J. R. (2016). Children's racial categorization in context. *Child Development*

Perspectives, 10(1), 33–38. https://doi.org/10.1111/cdep.12155

Piccinelli, M., & Wilkinson, G. (2000). Gender differences in depression: Critical review. *British Journal of Psychiatry, 177*(6), 486–492. https://doi.org/10.1192/bjp.177.6.486

Pilgrim, D. (2014). Historical resonances of the DSM-5 dispute: American exceptionalism or Eurocentrism? *History of the Human Sciences, 27*(2), 97–117. https://doi.org/10.1177/0952695114527998

Poland, J., & Caplan, P. J. (2004). The deep structure of bias in psychiatric diagnosis. In P. J. Caplan & L. Cosgrove (Eds.), *Bias in psychiatric diagnosis* (pp. 9–23). Lanham, MD: Jason Aronson.

Roche, B., & Barnes-Holmes, D. (2003). Behavior analysis and social constructionism: Some points of contact and departure. *The Behavior Analyst, 26*(2), 215–231. Retrieved from http://ezproxy.gsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2003-10336-003&site=ehost-live. Bryan.T.Roche@may.ie

Roche, B., Barnes-Holmes, D., Barnes-Holmes, Y., & Hayes, S. C. (2001). Social processes. In S. C. Hayes, D. Barnes-Holmes, & B. Roche (Eds.), *Relational frame theory: A post-Skinnerian account of human language and cognition* (pp. 197–209). New York, NY: Kluwer Academic/Plenum Publishers.

Roche, B., Barnes-Holmes, Y., Barnes-Holmes, D., Stewart, I., & O'Hora, D. (2002). Relational frame theory: A new paradigm for the analysis of social behavior. *The Behavior Analyst, 25*(1), 75–91. https://doi.org/10.1007/bf03392046

Rogers, W. A., & Mintzker, Y. (2016). Casting the net too wide on overdiagnosis: Benefits, burdens and non-harmful disease. *Journal of Medical Ethics, 42*, 717. https://doi.org/10.1136/medethics-2016-103715

Romans, S. E., Tyas, J., Cohen, M. M., & Silverstone, T. (2007). Gender differences in the symptoms of major depressive disorder. *The Journal of Nervous and Mental Disease, 195*(11), 905–911.

Schwartz, R. C., & Blankenship, D. M. (2014). Racial disparities in psychotic disorder diagnosis: A review of empirical literature. *World journal of Psychiatry, 4*(4), 133.

Skinner, B. F. (1957). *Verbal behavior*. East Norwalk, CT: Appleton Century Crofts.

Skinner, B. F. (1974). *About behaviorism*. Oxford, UK: Alfred A. Knopf.

Snowden, L. R. (2003). Bias in mental health assessment and intervention: Theory and evidence. *American Journal of Public Health, 93*(2), 239–243. https://doi.org/10.2105/ajph.93.2.239

Son Hing, L. S., Li, W., & Zanna, M. P. (2002). Inducing hypocrisy to reduce prejudicial responses among aversive racists. *Journal of Experimental Social Psychology, 38*(1), 71–78. https://doi.org/10.1006/jesp.2001.1484

Sue, S., Fujino, D. C., Hu, L.-T., Takeuchi, D. T., & Zane, N. W. (1991). Community mental health services for ethnic minority groups: A test of the cultural responsiveness hypothesis. *Journal of Consulting and Clinical Psychology, 59*(4), 533.

Sue, S., & Zane, N. (1987). The role of culture and cultural techniques in psychotherapy: A critique and reformulation. *American Psychologist, 42*(1), 37–45. https://doi.org/10.1037/0003-066x.42.1.37

Sue, S., Zane, N., Hall, G. C. N., & Berger, L. K. (2009). The case for cultural competency in psychotherapeutic interventions. *Annual Review of Psychology, 60*, 525–548. https://doi.org/10.1146/annurev.psych.60.110707.163651

Sukhera, J., & Watling, C. (2018). A framework for integrating implicit bias recognition into health professions education. *Academic Medicine, 93*(1), 35–40.

Trull, T. J., Jahng, S., Tomko, R. L., Wood, P. K., & Sher, K. J. (2010). Revised NESARC personality disorder diagnoses: Gender, prevalence, and comorbidity with substance dependence disorders. *Journal of Personality Disorders, 24*(4), 412–426.

Twohig, M. P., Domenech Rodriguez, M. M., & Enno, A. M. (2014). Promoting the multicultural comptency of psychological professionals though acceptance- and mindfulness-based methods. In A. Masuda (Ed.), *Mindfulness and acceptance in multicultural competency: A contextual approach to sociocultural diversity in theory and practice* (pp. 147–163). Oakland, CA: New Harbinger Publication.

Wakefield, J. C. (1999). The concept of disorder as a foundation for the DSM's theory-neutral nosology: Response to Follette and Houts, Part 2. *Behaviour Research and Therapy, 37*(10), 1001–1027.

Wakefield, J. C. (2007). The concept of mental disorder: Diagnostic implications of the harmful dysfunction analysis. *World Psychiatry, 6*(3), 149–156.

Watt, A., Keenan, M., Barnes, D., & Cairns, E. (1991). Social categorization and stimulus equivalence. *The Psychological Record, 41*(1), 33–50. Retrieved from http://ezproxy.gsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=1991-17495-001&site=ehost-live

Way, N., & Rogers, O. (2015). "[T] hey say Black men won't make it, but I know I'm gonna make it": Ethnic and racial identity development in the context of cultural stereotypes. In K. C. McLean & M. Syed (Eds.), *The Oxford handbook of identity development* (pp. 296–285). New York, NY: Oxford University Press.

Wegner, D. M. (1994). Ironic processes of mental control. *Psychological Review, 101*(1), 34–52.

Wenzlaff, R. M., & Wegner, D. M. (2000). Thought suppression. *Annual Review of Psychology, 51*(1), 59–91.

Whooley, O. (2014). Nosological reflections: The failure of DSM-5, the emergence of RDoC, and the decontextualization of mental distress. *Society and Mental Health, 4*(2), 92–110.

Widiger, T. A. (1998). Invited essay: Sex biases in the diagnosis of personality disorders. *Journal of Personality Disorders, 12*(2), 95–118.

Widiger, T. A., & Spitzer, R. L. (1991). Sex bias in the diagnosis of personality disorders: Conceptual and methodological issues. *Clinical Psychology Review, 11*(1), 1–22.

Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review, 107*(1), 101–126. https://doi.org/10.1037/0033-295x.107.1.101