Psychological Bulletin
1998, Vol. 123, No. 3, 260–278

Copyright 1998 by the American Psychological Association, Inc.
0033-2909/98/$3.00

Gender Differences in the Diagnosis of Mental Disorders: Conclusions and Controversies of the *DSM–IV*

Cynthia M. Hartung and Thomas A. Widiger
University of Kentucky

One of the more controversial issues in terms of mental disorder diagnoses has been their differential sex prevalence. The conclusions provided in the 4th edition of the American Psychiatric Association's (1994) *Diagnostic and Statistical Manual of Mental Disorders* were derived from systematic reviews of the research literature. However, this research is perhaps complicated by common sources of error. Two sources, in particular, are emphasized here: biases in sampling and biases within the diagnostic criteria themselves. The potential for such biases is illustrated for a wide variety of mental disorder diagnoses, and suggestions for research to address them are provided.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

One of the more controversial issues for mental disorder diagnoses has been their differential sex prevalence. The diagnoses that generated the most controversy in the development of the recent editions of the American Psychiatric Association's (APA's) influential *Diagnostic and Statistical Manual of Mental Disorders* (*DSM;* APA, 1980, 1987, 1994) were controversial largely because of their questionable application to one sex relative to the other (Ross, Frances, & Widiger, 1995). The critiques of decisions made for the third (*DSM–III*), revised third (*DSM–III–R*), and fourth (*DSM–IV*) editions of the *DSM* have often focused on issues concerning a purported sex bias (e.g., Brown, 1992; Caplan, 1991, 1995; Gallant & Hamilton, 1988; Kaplan, 1983b; L. E. A. Walker, 1994).

The statements provided in the *DSM–IV* concerning the differential prevalence of mental disorders for males and females represented authoritative conclusions of the APA derived from systematic reviews of the research literature (Frances, First, & Pincus, 1995). However, we indicate here that many of the conclusions regarding differential sex prevalence rates might be complicated by common sources of potential error or bias. An overview of these conclusions indicates a remarkable commonality of issues across diverse domains of clinical research. We emphasize two issues in particular: biases in sampling and biases within the diagnostic criteria themselves. These issues complicate the understanding of differential sex prevalence and of the nature of the disorders themselves.

Overview of the *DSM–IV*

Information provided in the *DSM–IV* on the differential sex prevalence for the 125 mental disorders described therein is presented in Tables 1–4. *DSM–IV* provides information regarding differential sex prevalence for 101 of these disorders (81%). No information is provided for the others, including dissociative amnesia, polysubstance-related disorders, hypoactive sexual desire disorder, dyspareunia, posttraumatic stress disorder (PTSD), reactive attachment disorder, and dementia due to HIV disease. The absence of information often reflected insufficient empirical research. However, the information provided in the *DSM–IV* was prepared by different individuals, without a published documentation of the bases for the conclusions. Therefore, judgments as to whether there was sufficient empirical support can vary substantially, and the bases for these judgments are at times unclear. For example, virtually no information regarding the sex ratio for PTSD is provided, despite the substantial amount of research on this disorder (Yonkers & Gurguis, 1995). Most epidemiological studies have reported a higher rate of PTSD in females (e.g., Breslau, Davis, Andreski, & Peterson, 1991; Helzer, Robins, & McEvoy, 1987), consistent with a higher rate of victimization (Yonkers & Gurguis, 1995), a lower resilience to stressors (Breslau et al., 1991), or both. A subtyping for the nature of the PTSD stressor would also yield substantial sex ratio variation (e.g., combat veterans vs. rape victims; Davidson & March, 1997). It is perhaps equally surprising that no information is provided for dissociative amnesia; dissociative disorders are diagnosed far more often in women (Spiegel, 1997), and dissociative amnesia is the diagnosis for alleged victims of childhood abuse, most of whom are women with purportedly recovered memories (Ross et al., 1995).

In contrast to the absence of information for PTSD and dissociative amnesia, the *DSM–IV* reports that factitious disorder occurs more often in males, despite an apparent lack of systematic research to support this conclusion. Factitious disorder occurs in two common variants that are gender related (Plewes & Fagan, 1994). One variant involves predominantly male wanderers with many hospitalizations across a wide geographic area; the female variant involves mostly younger patients with fewer hospitalizations and more single-symptom complaints. The male variant is often considered to be the classic form of the disorder, but it appears to occur with much less frequency than the female variant. In their comprehensive review of all cases of factitious

Cynthia M. Hartung and Thomas A. Widiger, Department of Psychology, University of Kentucky.

We would like to thank Richard Milich, Kenneth Sher, and William Grove for their review of earlier versions of this article, and John Ireland, Caryn Letcher, and Michelle Watson for their assistance in its preparation.

Correspondence concerning this article should be addressed to Cynthia M. Hartung, who is now at Department of Psychiatry, University of Chicago, 5841 South Maryland Avenue, MC 3077, Chicago, Illinois 60637. Electronic mail may be sent to chartung@yoda.bsd.uchicago.edu.

Table 1
Sex Ratios for Infancy, Childhood, and Adolescent Disorders

Disorder	Sex ratio
Mental retardation	1.5M:1F
Reading disorder	1.5–4M:1F
Mathematics disorder	—
Written expression disorder	—
Developmental coordination disorder	—
Expressive language disorder	M > F
Mixed language disorder	M > F
Phonological disorder	M > F
Stuttering	3M:1F
Autistic disorder	4–5M:1F
Rett's disorder	F only
Childhood disintegrative disorder	M > F
Asperger's disorder	M > F
ADHD	4–9M:1F
Conduct disorder	M > F
Oppositional defiant disorder	M > F
Pica	—
Rumination	M > F
Feeding disorder	M = F
Tourette's disorder	1.5–3M:1F
Chronic motor or vocal tic	—
Transient tic disorder	—
Encopresis	M > F
Enuresis	M > F
Separation anxiety disorder	F > M
Selective mutism	F > M
Reactive attachment disorder	—
Stereotypic movement disorder[a]	M > F

Note. Dashes indicate that no information was provided. M = male; F = female; ADHD = attention deficit hyperactivity disorder.
[a] Stereotypic movement disorder with head banging is said to be more prevalent in boys at a ratio of 3:1, but self-biting is said to be more common in girls.

disorder by proxy reported in the clinical literature (conducted for the *DSM–IV*), Plewes and Fagan (1994) indicated that 99% of patients were female and that half of them met the criteria for factitious disorder.

Sources of Potential Error

Accurate estimates of differential sex prevalence are difficult to obtain and are subject to substantial dispute and controversy (Eagly, 1995; Widiger & Spitzer, 1991). Deviations from actual prevalence rates are typically attributed to some form of sex bias (e.g., Caplan, 1991, 1995; Kaplan, 1983b; L. E. A. Walker, 1994). In this article, we define a sex bias as a systematic deviation from an expected value that is associated with the sex of the individual (Widiger & Spitzer, 1991). We are not suggesting that these deviations from an expected value "reflect society's sexism" (Kaplan, 1983a, p. 802), a "deeply entrenched sexism in the American Psychiatric Association" (Caplan, 1995, p. 169), or sexist attitudes on the part of researchers or authors of the *DSM–IV*. Systematic deviations from an expected value can result from well-intentioned, conscientious efforts to provide accurate estimates of differential sex prevalence rates.

A sex bias will usually represent a deviation from the true value, although in many instances the value that would be obtained from an unbiased sampling or unbiased assessment will not, in fact, be the most accurate value, as a result of other (earlier) biases in the process of diagnosis (Widiger & Spitzer, 1991). Ideally, the differential sex prevalence rates provided in the *DSM–IV* would reflect sex-based differences in biogenetic or environmental factors that contributed to the etiology (development, escalation, or course) of the disorder. For example, there appear to be male–female genetic and hormonal differences that contribute (at least in part) to the higher rate of

Table 2
Sex Ratios for Delirium, Dementia, Amnestic, and Substance-Related Disorders

Delirium, dementia, or amnestic disorder		Substance-related disorder[a]	
Disorder	Sex ratio	Disorder	Sex ratio
Delirium[b]	F > M	Alcohol	M > F
Dementia of Alzheimer's[b]	F > M	Amphetamine	M > F
Vascular dementia[b]	M > F	Caffeine	M > F
Dementia HIV disease[b]	—	Cannabis	M > F
Dementia head trauma	M > F	Cocaine[c]	M = F
Dementia Parkinson's[b]	—	Hallucinogen	3M:1F
Dementia Huntington's[b]	M = F	Inhalant	3–4M:1F
Dementia Pick's disease[b]	—	Nicotine	M > F
Dementia Creutzfeldt–Jakob[b]	—	Opioid	3–4M:1F
Amnestic disorders	—	Phencyclidine	2M:1F
		Sedative, hypnotic, anxiolytic	F > M
		Polysubstance	—

Note. Dashes indicate that no information was provided. M = male; F = female.
[a] Substance-induced disorders are not included in these tables because differential sex prevalence depends primarily on the respective substance, and findings for abuse, dependence, intoxication, and withdrawal are collapsed.
[b] Differential sex prevalence depends primarily on associated medical disorder. It is for this reason that other mental disorders (e.g., mood disorder) due to a general medical condition are not included in tables.
[c] "Unlike most other Substance-Related Disorders, with which males are more commonly affected than females, Cocaine Use Disorders are almost equally distributed between males and females" (American Psychiatric Association, 1994, p. 228).

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Table 3
Sex Ratios for Schizophrenic, Mood, Anxiety, Adjustment, Somatoform, Dissociative, and Sleep Disorders

Disorder	Sex ratio
Schizophrenia	M = F
Schizophreniform disorder	—
Schizoaffective disorder	F > M
Delusional disorder	M = F
Brief psychotic disorder	—
Shared psychotic disorder	F > M
Major depressive disorder	2F:1M
Dysthymic disorder	2–3F:1M
Bipolar I disorder	M = F
Bipolar II disorder	F > M
Cyclothymic disorder	M = F
Panic without agoraphobia	2F:1M
Panic with agoraphobia	3F:1M
Agoraphobia without panic	F > M
Specific phobia	F > M
Social phobia	F > M
Obsessive–compulsive disorder	M = F
Posttraumatic stress disorder	—
Acute stress disorder	—
Generalized anxiety	2F:1M
Adjustment disorder[a]	M = F
Somatization disorder	F > M
Undifferentiated somatoform disorder	F > M
Conversion disorder	2–10F:1M
Pain disorder	F > M
Hypochondriasis	M = F
Body dysmorphoric disorder	M = F
Factitious disorder	M > F
Dissociative amnesia	—
Dissociative fugue	—
Dissociative identity disorder	3–9F:1M
Depersonalization	—
Anorexia nervosa	>9F:1M
Bulimia nervosa	>9F:1M
Primary insomnia	F > M
Primary hypersomnia	3M:1F
Narcolepsy	M = F
Breathing-related sleep disorder	8M:1F
Circadian rhythm sleep disorder	—
Nightmare disorder	2–4F:1M
Sleep terror disorder	M = F
Sleepwalking disorder	M = F

Note. Dashes indicate that no information was provided. M = male; F = female.
[a] No information was provided for depressed mood, anxiety, or conduct subtypes, although substantial differential sex ratios are likely to be found for these variants.

rapid cycling in bipolar disorder in women (Leibenluft, 1996), differences in childhood traumatic experiences that contribute (at least in part) to a higher rate of dissociative disorders in women (Spiegel, 1997), and a variety of interactive biogenetic and environmental factors that result in the substantial preponderance of boys with conduct disorder (Eme & Kavanaugh, 1995). We would not consider such sex-related differences in the etiology of a disorder to represent a sex bias (although, as we indicate subsequently, it is at times suggested in the *DSM–IV* that comparable differences in early socialization do not contribute to actual differences in the differential sex prevalence of a disorder). The intent of this review of potential biases and errors in the estimates of differential sex prevalence rates is not to suggest that no differential sex prevalence rates actually exist. In fact, it is unlikely that the prevalence rate for any mental disorder will be virtually identical across the two sexes (Meehl, 1967), and the differences that do occur are likely to be significant and at times even substantial for many of the mental disorders (Anthony, Arria, & Johnson, 1995; Corbitt & Widiger, 1995; Eme & Kavanaugh, 1995; Gaub & Carlson, in press; Nolen-Hoeksema, 1995; Wolk & Weissman, 1995).

However, the precise value of the true differential sex prevalence rate will be unknown in most instances, and the findings that have been obtained will continue to be subject to charges of bias (e.g., Brown, 1992; Caplan, 1991, 1995; Kaplan, 1983a, 1983b; L. E. A. Walker, 1994; Worell & Robinson, 1995; Zoc-

Table 4
Sex Ratios for Sexual, Gender Identity, Impulse Control, and Personality Disorders

Disorder	Sex ratio
Hypoactive sexual desire disorder[a]	—
Sexual aversion disorder	—
Female sexual arousal	F only
Male erectile disorder	M only
Female orgasmic disorder[a]	F only
Male orgasmic disorder	M only
Premature ejaculation	M only
Dyspareunia[b]	—
Vaginismus	F only
Exhibitionism	M > F
Fetishism	M > F
Frotteurism	M > F
Pedophilia	M > F
Sexual masochism	20M:1F
Sexual sadism	M > F
Transvestic fetishism	M > F
Voyeurism	M > F
Gender identity disorder	2–3M:1F
Intermittent explosive disorder	M > F
Kleptomania	F > M
Pyromania	M > F
Pathological gambling disorder	2M:1F
Trichotillomania	F > M
Paranoid personality disorder	M > F
Schizoid personality disorder	M > F
Schizotypal personality disorder	M > F
Antisocial personality disorder	M > F
Borderline personality disorder	3F:1M
Histrionic personality disorder	F > M
Narcissistic personality disorder	1–3M:1F
Avoidant personality disorder	M = F
Dependent personality disorder	F > M
Compulsive personality disorder	2M:1F

Note. Dashes indicate that no information was provided. M = male; F = female.
[a] It was stated in the revised third edition of the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM–III–R*) that "Hypoactive Sexual Desire Disorder and Inhibited Orgasm are more common in females" (American Psychiatric Association [APA], 1987, p. 292), but no information regarding sex ratio was provided in the fourth edition (APA, 1994).
[b] It was stated in the *DSM–III–R* that "although Dyspareunia is defined so that it can occur in males, it rarely does" (APA, 1987, p. 292), but no information regarding sex ratio was provided in the fourth edition (APA, 1994).

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

colillo, 1993) until the conceptual and methodological issues discussed subsequently are adequately resolved. Two forms of sex bias considered here are biases due to sampling procedures and biases within diagnostic criteria. These sources of error are particularly important to consider because they appear to be common within diagnostic research and are interactive in their effects. Biases in sampling can contribute to the development of biased diagnostic criteria sets, and vice versa. Not considered in this article are biases in the process by which the diagnostic criteria are applied, including biases within assessment instruments (e.g., Lindsay & Widiger, 1995) and biases within clinicians (e.g., Potts, Burnam, & Wells, 1991). Biases within clinicians and assessment instruments are of substantial clinical and theoretical importance but are beyond the scope of our review. This form of sex bias has been addressed in many other recent reviews (e.g., Garb, 1997; Worell & Robinson, 1995).

Biases in Sampling

Most studies of mental disorders use nonprobability samples, obtained largely by convenience, that are highly vulnerable to selection biases (Sher & Trull, 1996). "Despite the importance of sampling in psychopathology research, there are surprisingly few general discussions of the issues and implications related to sampling choices" (Sher & Trull, 1996, p. 377). One common form of bias in sampling in psychological research has been a disproportionate representation of the sexes (Gannon, Luchetta, Rhodes, Pardie, & Segrist, 1992). Disproportionate representation of the sexes is common within clinical research, but it is also a substantial problem in the development of unbiased diagnostic criteria (National Institutes of Health [NIH], 1994).

Biased representation within clinical settings. Differential sex prevalence rates within clinical settings can reflect many male–female differences that are independent of sex-related differences in the etiology of a disorder, including (but not limited to) sex-related differences in the willingness to seek treatment, the ability or willingness to acknowledge the disorder's presence, the reactions of others to the syndrome, and the presence of comorbid disorders or conditions that affect the likelihood of seeking or receiving treatment. Allusions to such complicating factors are present in the discussion of differential sex prevalence rates throughout the *DSM–IV*, particularly in reference to discrepancies in the rates across clinical and community settings, as in the cases of conduct disorder, dependent personality disorder, histrionic personality disorder, separation anxiety disorder, schizophrenia, cyclothymic disorder, social phobia, nightmare disorder, generalized anxiety disorder, and pathological gambling (APA, 1994).

In some instances, it is suggested that the differential sex prevalence rate is exaggerated in females. For example, "Trichotillomania appears to be much more common among females than among males" (APA, 1994, p. 619), but it is also noted that this finding may reflect "differential treatment seeking based on cultural or gender-based attitudes regarding appearance (e.g., acceptance of normative hair loss among males)" (p. 620) rather than a true sex ratio for the disorder. Similarly, no differences are noted for cyclothymia, in part because "women with Cyclothymic Disorder may be more likely to present for treatment than men" (APA, 1994, p. 364). However, in other instances, it is suggested that the differential prevalence rate within clinical settings disproportionately favors males. For example, approximately two thirds of individuals with pathological gambling disorder are said to be male, but it is also indicated that "females are underrepresented in treatment programs for gambling and represent only 2%–4% of the population of Gamblers Anonymous" (APA, 1994, p. 616). "This may be a function of the greater stigma attached to female gamblers" (APA, 1994, p. 616). Similarly, "epidemiological and community-based studies suggest that Social Phobia is more common in women than in men. In most clinical samples, however, the sexes are either equally represented or the majority are male" (APA, 1994, p. 414). Likewise, "hospital-based studies suggest a higher rate of Schizophrenia in males, whereas community-based surveys have mostly suggested an equal sex ratio" (APA, 1994, p. 282).

One general point of distinction not noted in the *DSM–IV* is the differential sex prevalence rate of disorders of childhood versus adulthood that may reflect, at least in part, differences in the behaviors or symptoms that are recognized as being disordered in childhood versus adulthood. Of the 21 disorders usually first diagnosed in infancy, childhood, or adolescence for which sex ratios are provided, 17 are said to be more common in boys than in girls, 1 is said to be equally common in both sexes (feeding disorder), and only 3 are said to be more common in girls than in boys (see Table 1). In addition, many of the other childhood disorders for which no information is provided in the *DSM–IV* may also be more common in boys, such as tic disorders and disorders of written expression (Eme, 1992; Popper & Steingard, 1994). In contrast, of the 80 disorders diagnosed in adulthood for which sex ratios are provided, 35 are said to be more common in men than in women (17 of which are substance related or a paraphilia), 31 are said to be more common in women than in men, and 14 are said to be equally common in both sexes (see Tables 2–4). Thus, it appears that childhood mental disorders are overwhelmingly more common in boys than in girls, whereas, in adulthood, mental disorders are said to be more evenly distributed.

One might expect the ratio to be consistent from childhood to adulthood if there is a continuation of disorders across time. However, there does appear to be some inconsistency from childhood to adulthood, and this inconsistency may not be entirely due to the actual differential prevalence rates in children versus adults. For example, some of the differential sex prevalence rate may reflect, at least in part, the source of referral for treatment. In childhood, the motivation for treatment comes largely from other individuals; very few children initiate their psychiatric treatment. Treatment for mental disorders is typically provided to children at the request of others, notably parents and teachers (Goodman et al., 1997; Popper & Steingard, 1994). Consequently, the concerns presented for treatment may often reflect symptoms or behaviors that are especially troubling to a parent or a teacher (Shaywitz, Shaywitz, Fletcher, & Escobar, 1990). In contrast, the source for the initiation of treatment in adulthood is more likely to be the identified patient, with perhaps more adult women than men being motivated to seek treatment within clinical settings (Good & Wood, 1995). Major depressive disorder, dysthymic disorder, and trichotillomania are all reported in

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

the *DSM–IV* to be diagnosed equally in boys and girls during childhood but much more frequently in women during adulthood; sleep terror disorder likewise changes from a greater frequency in boys during childhood to an equal male–female frequency in adulthood.

As a result, the adult sections of the *DSM–IV* may be relatively more sensitive to disorders that are troubling to the identified patient, whereas the childhood sections may be relatively more sensitive to disorders that are troubling to others. Many of the disorders in the childhood sections are indeed externalizing disorders, or disorders that would be particularly troubling to parents (e.g., enuresis) or teachers (e.g., reading disorder), whereas the adult sections place relatively more emphasis on internalizing disorders (e.g., major depressive disorder and generalized anxiety disorder) that are particularly troubling to the identified patient. The most common childhood mental diagnoses concern disruptive behaviors (Popper & Steingard, 1994), three of which (i.e., attention-deficit hyperactivity disorder [ADHD], conduct disorder, and oppositional defiant disorder) are placed in a section titled (in part) "Disruptive Behavior Disorders" (APA, 1994, pp. 78–94). There is no comparable section for this domain of psychopathology within adulthood.

Sex differences in the prevalence rate of particular forms of child psychopathology were recently reviewed by Keenan and Shaw (1997). They indicated that before the age at which children begin school, there appear to be few sex differences in rates of difficult temperament, activity level, or noncompliance; substantial differences emerge in childhood; during adolescence, the pattern shifts to a female predominance; and, by adulthood, women are twice as likely to be suffering from the internalizing disorders of anxiety and depression. One explanation for this shift is a change in girls' early problem behaviors resulting from more rapid biological, cognitive, and social emotional development. However, the authors also provided an equally compelling explanation that "girls who were difficult as infants and toddlers are socialized to express their early problem behavior in a sex-stereotyped form (anxiety)" (Keenan & Shaw, 1997, p. 96). "Through the process of socialization, girls' early problem behavior is channeled into predominantly internalizing problems during the school-age period, which go either unnoticed or unidentified by parents and teachers" (Keenan & Shaw, 1997, p. 96), who are relatively more concerned with the externalizing disorders that are more evident in boys. By adulthood, there is a relative shift back toward an increased emphasis within clinical settings on the internalizing disorders, with the externalizing disorders receiving relatively less attention.

Many of the disorders of childhood showing male predominance are treated much less frequently in adulthood. Overall, there is more continuity than discontinuity in the sex ratios from childhood to adulthood (Keenan & Shaw, 1997). However, many disorders of childhood that show male predominance lack any clear adult variants. For example, very few adults seek treatment for reading disorder or expressive language disorder. It is perhaps not surprising that some developmental disorders lack adult variants, but many do continue into adulthood. There is compelling empirical support for an adolescent and adult variant of ADHD (Barkley, Fisher, Edelbrock, & Smallish, 1990; Nadeau, 1995; Wender, 1995), but its inclusion in the *DSM–IV* was not without some opposition and controversy (Shaffer, 1994). The *DSM–III* reported that the adult variant of childhood oppositional defiant disorder was passive–aggressive personality disorder (APA, 1980), but the latter diagnosis lost its official recognition in the *DSM–IV*, being demoted to a criteria set needing further research (Frances et al., 1995). No adult variant of oppositional defiant disorder is suggested in the *DSM–IV*. It has been argued that the failure to include diagnoses for adults with disorders of dyscontrolled anger or aggression has reflected a masculine bias in the recognition of psychopathology or a denial of psychopathology that would be more common in men than in women (Caplan, 1995). Proposals for such diagnoses have, in fact, been made (e.g., paraphiliac rapism and sadistic personality disorder), and they were opposed in part because they might be used to mitigate the criminal responsibility of aggressive, assaultive males (Widiger, 1995).

Shaywitz et al. (1990) have suggested that the predominance of boys with reading disorder (see Table 1) is due in part to the presence of comorbid conduct disorder symptomatology. "School-identified samples are almost unavoidably subject to a referral bias and . . . reports of an increased prevalence of reading disability in boys may reflect this bias in ascertainment" (Shaywitz et al., 1990, p. 998). They found that systematic, structured assessments revealed little or no sex difference in the rate of reading disorder, but boys, as a result of the presence of comorbid disruptive behavior problems, were referred for treatment more often than girls. "Behavioral problems significantly differentiate reading-disabled children who are and are not identified by their schools" (Shaywitz et al., 1990, p. 1002).

Referral biases are certainly not the complete explanation for the differential sex prevalence rates. The differential sex prevalence rate in most instances, including this one, will reflect gender-related etiological differences (Liederman & Flannery, 1995; Shaywitz et al., 1995). Nevertheless, these actual sex differences will often be magnified, confused, or masked within clinical settings by a variety of confounding factors, including, for example, comorbid conditions with a conflicting or complementary differential sex prevalence rate. Gaub and Carlson (in press) concluded, on the basis of their meta-analytic review of research examining gender differences in ADHD, that boys with ADHD show higher rates of comorbid conduct disorder and other disruptive behaviors that result in a significantly higher likelihood of clinic referral. They indicated that clinic-referred ADHD girls are not representative of ADHD girls in general and that studying gender differences within clinic-referred samples can contribute to erroneous conclusions regarding the nature of ADHD in girls. Barkley (1995a), summarizing the results of a conference on gender differences in ADHD sponsored by the National Institute of Mental Health (NIMH), indicated as well that "females sent to clinics may need to show more severe symptoms, relative to other females, before triggering such a referral" (p. 2). "A referral bias operates in determining which gender is likely to get referred for clinical services. This implies that females may require a more severe degree of ADHD and associated disruptive behaviors . . . than do males before such a referral is made" (Barkley, 1995a, p. 4).

It is suggested in the *DSM–IV* that there might, in fact, be an equal prevalence of males and females with dependent personality disorder, because the increased rate of females being diagnosed with the disorder (Bornstein, 1996, 1997) is "not

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

significantly greater than the sex ratio of females within the respective clinical setting" (APA, 1994, p. 667). This conclusion was based largely on a study of the differential sex prevalence of personality disorders conducted by Reich (1987; see also Hirschfeld, Shea, & Weise, 1991). Reich indicated that 75% of the individuals with dependent personality disorder in his sample were female, but he concluded that this did not represent a greater prevalence of females than males with the disorder in that the rate was not significantly higher than the 63% of females within his sample of patients. "Previous estimates may not have taken the underlying excess of women in the psychiatric population into account" (Reich, 1987, p. 487).

Correcting for the prevalence of females within a clinical setting when estimating the prevalence for a particular disorder might appear responsive to normative base rate expectations (Widiger, Hurt, Frances, Clarkin, & Gilmore, 1984), but it would not be appropriate when the disorder being researched is the major reason for the sex ratio obtained at the clinic (Corbitt & Widiger, 1995). One is then correcting for the presence of the disorder that is being researched. For example, one would not conclude that males are just as likely to have bulimia as females because the proportion of bulimics who were female was not significantly higher than the base rate of females at an eating disorders clinic. The reason that most of the patients at the clinic were female might be, in large part, that most persons with bulimia are female. In Reich's (1987) sample, 47% of the sample had a dependent personality disorder. Requiring that the proportion of dependent females be greater than the proportion of females within the sample to conclude that more females than males have the disorder is perhaps comparable to requiring that the proportion of females with bulimia be greater than the proportion of females with an eating disorder.

Nevertheless, correcting for the female base rate within a clinical setting to determine the actual differential sex prevalence rate is appropriate when the women within the clinic are there (or men are not there) for reasons other than the disorder being researched. For example, concluding that there are more men than women with dependent personality disorder on the basis of research within a Veterans Administration hospital would clearly be inaccurate. More men than women will have personality disorders within a Veterans Administration hospital because there are more male than female veterans. However, the reasons for the differential base rate within most clinical settings will usually be unclear, particularly when one considers the additional complexity of the overlap, co-occurrence, and comorbidity among different mental disorders. For example, in Reich's (1987) study, 52% of the sample (n = 88) was obtained randomly from new admissions to an outpatient clinic, and 48% (n = 82) was obtained from an ongoing study of panic disorder. Panic disorder is said to occur two to three times more often in females than in males (APA, 1994). Therefore, Reich's finding that 75% of the individuals with dependent personality disorder were female could have been due to the disproportionate sampling of persons with a panic disorder. However, it could also be that most of the cases of dependent personality disorder were obtained from the randomly selected new admissions, or the presence of individuals with panic disorder seeking treatment might have been due at least in part to the presence of dependent personality traits (Phillips & Gunderson, 1994).

Biased representation within empirical studies. If a clinical setting fails to have a representative sample of the individuals with a disorder, then studies that sample randomly from that setting will be comparably biased. Disproportionate representation of the sexes within clinical studies may, in fact, be the rule rather than the exception.

For example, research on conduct disorder has been confined largely to boys (Goodman & Kohlsdorf, 1994; Robins, 1986; Zahn-Waxler, 1993; Zoccolillo, 1993). Confining such research largely to boys does have a compelling rationale; conduct disorder appears to occur much more often in boys than in girls (APA, 1994; Eme & Kavanaugh, 1995). Researchers who wish to study prototypic cases might then confine their study to boys, because the inclusion of girls would increase the heterogeneity of the sample and perhaps complicate analyses and interpretation as a result of the additional gender variance (Zoccolillo, 1993).

However, confining research to boys can also contribute to an inaccurate (e.g., male-biased) description of the disorder (Goodman & Kohlsdorf, 1994) that can, in turn, affect the future identification (i.e., diagnosis) of the disorder, contributing to further disproportionate sampling of the sexes (Robins, 1986, 1991). For example, Zoccolillo (1993) suggested that a childhood onset for conduct disorder is a description of a male variant of the disorder, because the onset may occur later in girls. Childhood versus adolescent onset of conduct disorder does have substantial empirical support, but it may have relatively more validity or significance for boys (Webster-Stratton, 1996), at least as it has been researched to this point (e.g., Moffitt, Caspi, Dickson, Silva, & Stanton, 1996). Robins (1986, 1991) has suggested likewise that the diagnostic criteria that have been developed for conduct disorder have been biased in favor of a male variant. She argued that the *DSM–III–R* provided a major shift in the description of the disorder, replacing many nonaggressive acts (e.g., early sexual activity and substance use) with quite severe aggressive behavior (e.g., cruelty to animals, physical cruelty to people, and sexual assault or rape) in large part as a result of the research being predominated by the way in which the disorder presents itself in boys.

Similar concerns have been raised for many other mental disorders, including antisocial personality disorder (Rutherford, Alterman, Cacciola, & Snider, 1995), ADHD (Loeber & Keenan, 1994; Popper & Steingard, 1994), alcohol abuse (Brett, Graham, & Smythe, 1995), schizophrenia (E. F. Walker & Lewine, 1993), and anxiety disorder (Blanchard, Griebel, & Blanchard, 1995). For example, 254 empirical studies were published in the *Journal of Abnormal Child Psychology* in the 6-year period from the appearance of the *DSM–III–R* in 1987 to the appearance of the *DSM–IV* in 1994. Eleven concerned adults or did not specify the sex of the participants. Of the remaining 243, 71% (173) included both boys and girls, whereas 29% were confined to one sex. Of the 70 studies confined to one sex, 99.6% (69) were studies of male children; only 1 was a study of female children (i.e., Nangle & Foster, 1992). Seventy of the 243 empirical studies concerned ADHD. Only half of them (49%) sampled both male and female children, and 100% of the remaining 36 studies were confined to boys. A total of 4,873 children participated in these 70 studies of ADHD.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Eighty-one percent of them (3,967) were boys, and only 19% (906) were girls.

Confining ADHD research largely to boys does, again, have a compelling rationale; ADHD (the predominantly hyperactive–impulsive type) appears to indeed occur much more often in boys than in girls (Barkley, 1990, 1995a; Gaub & Carlson, in press). However, deriving a set of diagnostic criteria for ADHD based on this research could provide excessive weight to how the disorder is expressed in boys. "ADHD might at present be defined by male-specific characteristics, with a deemphasis of the attentional or affective deficits seen in young females" (Popper & Steingard, 1994, p. 825). Some researchers suggest that there are important differences in how the disorder appears in boys and girls. Gaub and Carlson (in press) concluded, on the basis of their meta-analytic review of research on gender differences in ADHD, that girls with ADHD tend to display greater intellectual impairment, lower levels of hyperactivity, and lower rates of externalizing behaviors; however, the authors also emphasized that their findings were limited by concerns regarding gender differences in clinical versus community settings. Barkley (1995a), summarizing the results from the NIMH conference on gender differences in ADHD, concluded that, within community samples, females show lower ratings of inattention, less hyperactivity on ratings, fewer impulsive errors on continuous performance tests, and faster reaction times. He indicated, however, that such differences "are often found in normal samples and they do not reflect any differences that would be specific to the disorder or abnormal" (Barkley, 1995a, p. 1). "In clinic referred samples, many gender differences in ADHD symptoms disappeared, such that females appear to differ from males only in demonstrating less hyperactivity" (Barkley, 1995a, p. 1). Barkley (1995a) concluded that, overall, "there is no evidence of a qualitatively different expression of the disorder in males than in females" (p. 4). Nevertheless, he indicated that there were sufficient quantitative differences such "that girls would need a lower cutoff score than boys if one were attempting to apply the same threshold for deviance. . . . In other words, a cutoff score should be gender-based with girls" (Barkley, 1995a, p. 2).

The substantial overlap and confusion of ADHD with conduct disorder may itself reflect, in part, the impact of gender variables. Loeber and Keenan (1994) concluded, on the basis of their review of ADHD comorbidity research, that there "appears to be a fundamental difference in the nature of comorbid conditions as they emerge over time in females and males. The importance of this gender effect, however, has not been sufficiently recognized by researchers and clinicians" (Loeber & Keenan, 1994, p. 517). Gaub and Carlson (in press) likewise suggested that the predominance of externalizing symptomatology in boys with ADHD may reflect, at least in part, this comorbid symptomatology. Eme and Kavanaugh (1995) made the complementary point for the comorbidity of conduct disorder with ADHD, indicating that "the marked sex difference in ADHD doubtlessly makes a major contribution to explaining the marked sex difference in Conduct Disorder" (p. 408).

The *DSM–IV* criteria for ADHD have broadened substantially, particularly with respect to an increased recognition of symptoms of inattention without either hyperactivity or impulsivity (Frances et al., 1995). Lahey et al. (1994) reported, on the basis of the *DSM–IV* ADHD field trial, that "the broader DSM-IV definition resulted in the diagnosis of previously unidentified youths at a 4:1 ratio" (p. 1682). More important for the present purposes, the *DSM–IV* revisions may also change significantly the sex ratio of the disorder. "The new cases identified . . . by DSM-IV are more than twice as likely to be female . . . with most of the newly identified girls being in the predominantly inattentive type" (Lahey et al., 1994, p. 1682), consistent with the gender differences noted in prior research on ADHD (Gaub & Carlson, in press).

Concerns regarding potential sampling biases in alcohol abuse and dependence research have been present for some time (e.g., Emrick, 1974; Vannicelli & Nash, 1984). Brett et al. (1995) considered every empirical study in the 17 leading English-language journals concerned with sustance abuse or dependence published during the year 1992. Of the 1,041 articles, 601 were empirical studies with human participants. Only 6.3% included predominantly female participants (i.e., at least 80% female), whereas 30.3% were predominantly male. In 73.3% of the studies, the majority of the participants were male. Brett et al. also indicated a tendency of the authors to highlight the presence of females as a qualification to the findings, but not the predominance of males. "Research conducted on male samples was often described as if male experience was generalizable to both males and females" (Brett et al., 1995, p. 32). The National Institute on Alcohol Abuse and Alcoholism has emphasized for some time a need to limit references to a general alcohol abuse or alcohol dependence to findings that pertain equally to both sexes (Blume, 1990); this has been neglected in most studies, however, despite the substantial data indicating that the phenomenology, course, and treatment of alcohol abuse and dependence vary significantly across genders (Schmidt, Klee, & Ames, 1990; Toneatto, Sobell, & Sobell, 1992). As a result, alcohol abuse and dependence, as described in the scientific and clinical literature, might be to some extent a male alcohol abuse and dependence, not alcohol abuse and dependence in general (Hughes, 1990; Wilke, 1994).

A final illustration is provided by research on schizophrenia. Schizophrenia is a disorder for which "it has long been held that males and females are affected in roughly equal numbers" (APA, 1994, p. 281). One would then expect there to be roughly an equal number of male and female participants in schizophrenia research, but this has not been the case. Wahl (1977) had reported, before publication of the *DSM–III,* that male participants outnumbered female participants in schizophrenia research by almost two to one over a 2.5-year period in a systematic review of the *Journal of Abnormal Psychology, Archives of General Psychiatry,* and the *Journal of Nervous and Mental Disease* from 1974 to 1976. Wahl and Hunter (1992) indicated, more recently, that little has changed. They considered every study on schizophrenia published in *Archives of General Psychiatry,* the *American Journal of Psychiatry,* the *Journal of Abnormal Psychology,* and the *Journal of Nervous and Mental Disease* over a 5-year period from 1985 through 1989. One hundred ninety-eight studies were identified. Approximately 75% included more males than females; only 8% had more females than males. Forty-nine of the studies (25%) were confined to males, and only one was confined to females. Summing across all studies, 5,180 of the individuals with schizophrenia who

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

were researched were male (69%), whereas 2,311 were female (31%).

Sex differences in schizophrenia have always been of substantial theoretical and clinical interest, leading to the recognition of differences in onset, premorbid history, course, marital functioning, treatment responsivity, and, possibly, brain abnormalities (Goldstein & Tsuang, 1990). "With the accumulating evidence that male and female schizophrenic individuals are not alike, it is clear that the axiom of analyses of research by gender must apply" (Wahl & Hunter, 1992, p. 313). Nevertheless, such analyses appear to be the exception rather than the rule. "The end result is a significant overrepresentation of males in schizophrenia research—a bias that has persisted for at least 15 years—and knowledge of schizophrenia that is really knowledge of male schizophrenia" (Wahl & Hunter, 1992, p. 316).

Gender sampling biases might, in fact, explain some of the sex differences that have been obtained. E. F. Walker and Lewine (1993) have argued that the purportedly better treatment responsivity of females with schizophrenia may reflect in part gender differences in the recognition of the presence of a psychopathology or need for treatment. Goldstein (1993) disagreed with their conclusions but did agree that a "fruitful approach to understanding discrepancies in studies of gender differences may be to fully characterize the disorder within males and females as well as between the sexes. . . . [Yet,] in numerous studies, even the distribution of the sample by sex is unreported" (p. 13).

Biases Within Diagnostic Criteria

Concerns regarding gender-biased sampling can be addressed in part through an increased use of probability (epidemiologic) sampling of the population, proportionate representation of the sexes within studies, and analysis of findings by gender. These recommendations are discussed in more detail later. However, an inaccurate estimate of the differential sex prevalence of a disorder will still 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.

Potential gender biases within diagnostic criteria have received substantial attention in the clinical literature (Brown, 1992; Caplan, 1995; Kaplan, 1983b; Landrine, 1989; Wakefield, 1987; L. E. A. Walker, 1994). Ideally, diagnostic criteria sets would be gender neutral. However, many of the criteria sets may disproportionately favor the manner in which the disorder appears in one gender relative to the other. Gender neutrality is particularly difficult to achieve for disorders that are expressed or appear differently in males and females or involve, at least in part, maladaptive variants of gender-related behaviors. As suggested in a study of 2,013 adolescents conducted by Huselid and Cooper (1994), this may include a wide variety of disorders.

> Our findings replicated gender differences in patterns of symptom expression, with female adolescents reporting more psychological distress and lower self-esteem and male adolescents reporting more delinquency and substance use. More important, these results reveal that gender roles account for a substantial portion of the sex differences in both internalizing and externalizing symptoms. (Huselid & Cooper, 1994, p. 600).

Implications of this issue for diagnostic criteria sets are illustrated here with respect to orgasmic and sexual desire disorders, histrionic personality disorder, conduct disorder, and somatization disorder.

Orgasmic and sexual desire disorders. It was stated in the *DSM–III* that "inhibited sexual desire and inhibited orgasm are more common in females" (APA, 1980, p. 278). No statements regarding the differential sex prevalence of these disorders are provided in the *DSM–IV* (APA, 1994). Kaplan (1983b) had argued that the higher rate for women was due in part to "masculine-biased assumptions about what behaviors are healthy and what behaviors are crazy" (p. 786). "That the *DSM-III* . . . identifies Inhibited Orgasm and Inhibited Sexual Desire as (a) disorders and (b) disorders more commonly found in females, in the light of psychiatry's past mistake regarding women's sexuality, should encourage some thought when one is considering women's sexual pathology" (Kaplan, 1983b, p. 791).

Wakefield (1987), however, responded that the *DSM–III* criteria for inhibited orgasm were at least less sex biased than, for example, the prior Masters and Johnson (1982) criteria. Masters and Johnson diagnosed inhibited orgasm in women if they failed to achieve orgasms, whereas men had to be unable to have orgasms. In addition, women could be diagnosed with inhibited orgasms even if they never felt any arousal, whereas men had to experience a full arousal without an orgasm. In other words, the threshold for the diagnosis appeared to be much lower for women than for men, perhaps contributing to differential sex prevalence rates. The *DSM–III* criteria for inhibited orgasm were relatively less biased because they were more consistent across men and women. The only inconsistency in the *DSM–III* was the requirement for women that the clinician determine whether the sexual activity was "adequate in focus, intensity, and duration" (APA, 1980, p. 280) to lead to orgasm. However, "the fact that the female criterion specifies these additional conditions means that, at least in theory, [these criteria are] biased toward greater diagnosis of males" (Wakefield, 1987, p. 470).

The *DSM–IV* continues to provide somewhat different criteria for men and women with an orgasmic disorder. For a diagnosis in women, the clinician must consider whether their orgasmic capacity is less than would be reasonable for their sexual experience, and it is noted that women "exhibit wide variability in the type or intensity of stimulation that triggers orgasm" (APA, 1994, p. 506). No such qualifications are made for the diagnosis in men. Any difference in the criteria across men and women might suggest a potential bias (Wakefield, 1987). "Although there is a certain sense to the asymmetry, it is still on balance unwarranted and should be eliminated" (Wakefield, 1987, p. 470).

The different criteria in the *DSM–IV* for men and women are intended to be responsive to actual differences in physiology and prior experiences. The intention is to eliminate the effects of differences between men and women that would contribute to artifactual differences in the diagnosis of an orgasmic disorder (Schover, 1996). If less stimulation is necessary to trigger an orgasm in men, then perhaps more stimulation should be required before the diagnosis is made in women.

Wakefield (1987), however, emphasized that the purpose of different criteria for men and women should not be to obtain

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

an equal prevalence rate across men and women but to make the criteria equivalent. If there are differences in experiences or physiology that result in a different liability for or predisposition toward the development of the disorder, then the threshold of the diagnosis should not be adjusted to compensate for these actual differences in etiology. Therefore, if less stimulation is necessary to trigger an orgasm in men, it is reasonable to expect there to be less impairment in their ability to experience orgasms. Raising the threshold for the diagnosis in women to obtain comparable prevalence rates might in fact contribute to an underdiagnosis of the disorder in women.

It might be similarly misleading to alter the threshold for a diagnosis of an orgasmic disorder to correct for differences in the experiences of men and women, including experiences that are secondary to differences in socialization. A history of social cultural discouragement of sexual responsivity in women would contribute to less sexual responsivity in women. Yet, it is suggested in the *DSM–IV* that social cultural experiences that discourage sexual responsivity in women should be excluded as a contribution to the development of the disorder. "Clinical judgments about the presence of a Sexual Dysfunction should take into account the individual's . . . cultural . . . and social background, which may influence sexual desire, expectations, and attitudes. . . . For example, in some societies, sexual desires on the part of the female are given less relevance" (APA, 1994, p. 495).

Histrionic personality disorder. A diagnosis for which charges of sex-biased criteria have been raised frequently is histrionic personality disorder (Brown, 1992; Chodoff, 1982; Cooper, 1987; Kaplan, 1983b; Landrine, 1989; L. E. A. Walker, 1994). The diagnostic criteria for histrionic personality disorder include features that are related to stereotypically feminine behavior (e.g., emotional lability, concern with physical attractiveness, and sexual seductiveness; Sprock, Blashfield, & Smith, 1990). Males are less likely to meet the threshold for the diagnosis, and Kaplan (1983b) even suggested that "a healthy woman automatically earns the diagnosis of Histrionic Personality Disorder" (p. 789).

Others, however, have argued that a differential sex prevalence rate for a disorder that represents a maladaptive variant of stereotypically feminine traits is to be expected (Widiger & Spitzer, 1991). Research has suggested differences in the personalities of males and females (Eagly, 1995; Feingold, 1994) due in part to social cultural experiences and biogenetic differences (Buss & Schmitt, 1993; Hamilton & Jensvold, 1992). Comparable differences in the rate of personality disorders should then be expected (Corbitt & Widiger, 1995). For example, if females are, on average, more emotionally expressive than males (Feingold, 1994), then more females than males are likely to have a disorder characterized in part by dyscontrolled or excessive emotionality. Raising the threshold for the diagnosis of histrionic personality disorder in females would correct for these differences, but the differential sex prevalence rate might be a natural, expected result of biogenetic variables and social cultural experiences that vary across males and females.

However, any disorder that involves gender-related behaviors will be susceptible to sex biases within its diagnostic criteria. For example, one of the *DSM–III–R* criteria for histrionic personality disorder was being "overly concerned with physical attractiveness" (APA, 1987, p. 349). Kaplan (1983b) argued that the diagnostic criteria for histrionic personality disorder were sex biased (in part) because normal (healthy) women would meet the diagnostic criteria for the disorder. Widiger and Spitzer (1991) responded that this is unlikely because "the items describe maladaptive traits and normal false positive cases are not likely to occur. The person must be 'overly concerned' with physical attractiveness" (p. 14). Being overly concerned with physical attractiveness does, in fact, refer to a maladaptive behavior.

However, it is not at all clear when concern with physical attractiveness becomes a sufficiently maladaptive overconcern to warrant a diagnosis of a mental disorder. There is an apparent potential in this instance to confuse a nonmaladaptive adherence to social norms with a symptom of a mental disorder. To the extent that women are more concerned with or exert more time and effort toward their physical attractiveness (Sprock et al., 1990), misattributions of this disorder would occur more often for women than for men.

The authors of the histrionic diagnostic criteria for the *DSM–IV* revised this criterion to "consistently uses physical appearance to draw attention to self" (APA, 1994, p. 658). However, the intention of this revision was to construct a criterion that would be equally applicable to males and females (Pfohl, 1991). Both sexes might be equally likely to use their physical appearance to draw attention to themselves. To emphasize this point, it is noted in, the text of the *DSM–IV,* that "a man with this disorder may dress and behave in a manner often identified as 'macho' and may seek to be the center of attention by bragging about athletic skills" (APA, 1994, p. 656). Frances et al. (1995) suggested that the diagnosis of histrionic personality disorder has placed "too little emphasis on items and examples that would tap the parallel 'macho' male version expressing exaggerated masculine traits" (p. 373). They suggested that there may be as many histrionic males as histrionic females. "In his own way, Stanley Kowalski in *A Streetcar Named Desire* is just as histrionic as Blanche Du Bois" (Frances et al., 1995, p. 373).

The *DSM–IV* version may be even more problematic than the *DSM–III–R* version, because the diagnostic criterion no longer includes or requires that the behavior be excessive or maladaptive (Widiger, Mangine, Corbitt, Ellis, & Thomas, 1995). Consistently using one's physical appearance to draw attention to oneself is not inherently or necessarily maladaptive. "Wearing makeup every day might [be said to] constitute a consistent use of physical appearance to draw attention to oneself" (Widiger et al., 1995, p. 126) that would be closely associated with female gender and not associated with any psychopathology. Many normal women wear makeup every day to increase their physical attractiveness and thereby draw attention to themselves.

Conduct disorder. Robins (1991), an influential source for the *DSM–III* criteria for conduct disorder, suggested that the *DSM–III–R* was placing too much emphasis on the aggressive, violent behavior that is more evident within boys with conduct disorder and neglecting the antisocial behaviors that are more common in girls with conduct disorder, such as early use of alcohol and drugs, early sexual activity, violations of rules, and poor academic achievement. Robins and Price (1991) suggested lowering the threshold for the diagnosis in girls. Rutherford et

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

al. (1995) noted as well the predominance of items concerning aggression and violence within the *DSM–III–R* criteria for conduct disorder. They compared the *DSM–III* and *DSM–III–R* criteria in men and women and reported that "the six *DSM-III* childhood [conduct disorder] criteria that were dropped in *DSM-III-R* were the most frequently endorsed, were the most reliable, and had the highest item-total correlations of all childhood criteria for women" (Rutherford et al., 1995, p. 1315). They also suggested lowering the threshold for the diagnosis in women to compensate for an apparent bias within the diagnostic criteria. Zoccolillo (1993) has similarly recommended the development of separate diagnostic criteria for conduct disorder in girls to recognize that the disorder is expressed differently across boys and girls, placing relatively more emphasis on rule violations at home and school, substance abuse, prostitution, chronic lying, running away from home overnight, and poor school performance and less emphasis on vandalism, fire setting, burglary, use of a weapon in fights, stealing with confrontation of a victim, and rape.

Two new criteria were added to the *DSM–IV* diagnosis of conduct disorder: "often stays out at night despite parental prohibitions, beginning before age 13 years" and "often bullies, threatens, or intimidates others" (APA, 1994, p. 90). It is stated in the *DSM–IV* that the addition of these two symptoms "is based on the field-trial results and provides a definition that includes behaviors characteristic of females with Conduct Disorder" (APA, 1994, p. 775). However, in the report of the results from the *DSM–IV* disruptive behavior disorders field trial by Frick et al. (1994), staying out late at night was more predictive of conduct disorder in boys than in girls (positive predictive power values of .88 and .47, respectively) and was one of the two most predictive symptoms of the disorder in boys. The two most predictive symptoms in girls were initiating physical fights with others and being physically cruel to people (positive predictive power values of .92 and 1.00, respectively), symptoms with substantial aggressive content. They stated that the symptom that was "included as a test symptom based on its potential for predicting [conduct disorder] in girls" (Frick et al., 1994, p. 534) was ridiculing someone outside of the household in a mean way more than once, but this proposed criterion obtained very low predictive validity for both girls and boys. Nevertheless, staying out late at night might still prove to be useful for the diagnosis of conduct disorder in girls, given the findings of prior research (Goodman & Kohlsdorf, 1994; Robins, 1991; Zoccolillo, 1993).

It is also noted in the discussion of conduct disorder in the *DSM–IV* that "males with . . . Conduct Disorder frequently exhibit fighting, stealing, vandalism, and school discipline problems; [whereas] females with . . . Conduct Disorder are more likely to exhibit lying, truancy, running away, substance use, and prostitution" (APA, 1994, p. 88). To the extent that these are in fact valid differences in the way conduct disorder occurs in boys and girls, it might be unrealistic to attempt to use the same (i.e., gender-neutral) diagnostic criteria for boys and girls (Zoccolillo, 1993). The behaviors that are more specific to girls might be more diagnostic of the disorder in girls than any set of gender-neutral criteria, and the behaviors that are more specific to boys do appear to be somewhat biased against girls. For example, prostitution might be highly diagnostic of conduct disorder in girls, but it is not included within the *DSM–IV* diagnostic criteria set as a result of its lack of value in diagnosing conduct disorder in boys. On the other hand, forcing someone into a sexual activity is included within the diagnostic criteria set despite its lack of value in diagnosing the disorder in girls. The *DSM–IV* criteria set may not be the most valid criteria set for girls (or, perhaps, for boys), representing instead an unclear and inconsistent compromise of biases.

Zahn-Waxler (1993), however, has argued that any differences in the criteria sets for boys and girls would be problematic. The disorder may indeed appear differently across boys and girls, but alternative criteria sets will probably fail to provide the same threshold for the diagnosis in boys and girls (as suggested by Wakefield, 1987, for orgasmic disorder). Alternative criteria sets may instead define the disorder differently for boys and girls. For example, placing less emphasis on confrontational, violent, and aggressive behavior in the diagnosis of conduct disorder for girls, along with more emphasis on rule violations and deceitfulness, could have the effect of labeling (and stigmatizing) as mentally disordered in girls a level of behavioral dysfunction that is considered to be acceptable or normative when it occurs in boys. Boys would have to exhibit the more severe behavior of violent aggression to be diagnosed with the disorder, whereas girls would not. Zahn-Waxler argued that conduct disorder involves aggressive, confrontational, and violent behavior, and girls receive the diagnosis less often because they are less likely to engage in these behaviors. It may indeed be the case that girls who fall below the threshold for this diagnosis should receive a mental disorder diagnosis, but this disorder would not be conduct disorder. It would be a different or at least a less severe disorder that might also occur more frequently in boys.

Somatization disorder. Somatization disorder was originally diagnosed as hysteria (Guze, 1967), a diagnosis for which there have historically been substantial concerns regarding sex bias (Chodoff, 1982); the term *hysteria* translates literally as a wandering uterus. The name was changed to Briquet's syndrome (and, subsequently, to somatization disorder) in part to disassociate the diagnosis from any presumptions regarding its gender specificity. However, its criteria may continue to be biased heavily in favor of diagnosing the disorder predominantly in females (Wool & Barsky, 1994).

The *DSM–IV* criteria for somatization disorder were derived from the criteria set for Briquet's syndrome developed by Guze (1967). Guze required the presence of 20–25 of 59 symptoms obtained from at least 9 of 10 groups. In a comparison of the disorder in males and females, Guze, Woodruff, and Clayton (1972) concluded "that a diagnosis of hysteria, which is based upon a polysymptomatic disorder beginning early in life, is more likely to be made in women than in men" (p. 121). However, this was perhaps compelled in part by the diagnostic criteria; none of the 4 symptoms in Group 7 (i.e., dysmenorrhea, menstrual irregularity, amenorrhea, and excessive menstrual bleeding) could have occurred in men, nor would (or could) many of the symptoms in Group 8 (i.e., frigidity, vomiting during 9 months of pregnancy, sexual indifference, and dyspareunia). At least 1 symptom from 9 of the 10 groups was required for the diagnosis, but males were largely excluded from 2 of them. Guze et al. (1972) considered the inclusion of gender-specific

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

criteria appropriate because their male–female comparisons indicated that women expressed a significantly greater degree of somatization than men. "The men reported a mean 9.6 symptoms, and the women a mean of 13.4" (Guze et al., 1972, p. 123). However, the authors had not made any adjustments in this comparison for the inclusion of at least six gender-specific criteria (the only other male–female difference they reported was that the peak number of symptoms occurred at an earlier age in women).

Concerns regarding the apparent bias within the Guze et al. (1972) diagnostic criteria were raised in the clinical and research literature (e.g., Kaminsky & Slavney, 1976). However, the major concern for the *DSM–III* was to provide a simpler, more user-friendly criteria set, because the Guze (1967) algorithm of 20 of 59 symptoms from 9 of 10 groups was too cumbersome for general clinical use (Cloninger, 1996). The 59 symptoms were reduced to 37 on the basis of extensive analyses of 1,116 individuals sampled from five different clinical settings, none of whom were male (Cloninger, 1987).

It was evident to the authors of the *DSM–III* somatization disorder section that the criteria might not be as valid for males, given that one entire subset of items continued to be confined to "female reproductive symptoms" (APA, 1980, p. 243) such as painful menstruation, menstrual irregularity, excessive bleeding during menstruation, and severe vomiting during pregnancy. Therefore, only 12 of the 37 criteria were required in males, as compared with 14 in females. However, this lower threshold was selected arbitrarily. No data were obtained from males to assess whether the lower threshold was meaningful, valid, or successful in correcting for the apparent sex bias within the diagnostic criteria. Cloninger, Martin, Guze, and Clayton (1986) reported subsequently that the threshold for males would have to be lowered to 8 to equalize the prevalence, but it was unclear whether the goal of a reduced threshold was, in fact, to equalize the prevalence rate.

In the *DSM–III–R,* "the symptom list was revised so that the number of symptoms required for the disorder is the same for males and females" (APA, 1987, p. 423). However, no studies on males with somatization disorder informed this revision, the four female reproductive symptoms were retained, and only one male-biased item was added (i.e., impotence) that was perhaps offset by other items occurring more often in females (e.g., sexual indifference and pain during intercourse). One of the innovations of the *DSM–III–R* revision was to highlight an even briefer set of seven items that could be used as a screening device. These seven items had been identified by Othmer and DeSouza (1985) on the basis of studies of two independent samples of women with somatization disorder. Again, no males were included in the derivation of the criteria set. One of the seven screening criteria was dysmenorrhea (revised to painful menstruation in the *DSM–III–R;* APA, 1987). To help clinicians remember the seven criteria, Othmer and DeSouza (1985) provided a mnemonic formed by the first letter of each criterion: "somatization disorder besets ladies and vexes physicians" (p. 1148). This does suggest a degree of presumption regarding the sex that is to receive the diagnosis.

The revisions for the *DSM–IV* were substantial. The criteria are now organized into four subsections, with at least one symptom required from each area. However, one of these subsections again concerns female reproductive symptoms. In the 1991 draft of the *DSM–IV,* this subsection consisted of "a history of at least one sexual or reproductive symptom other than pain (such as sexual indifference, impotence, irregular menses, excessive bleeding, vomiting throughout pregnancy" (Task Force on *DSM-IV,* 1991, p. I:1). The potential bias against males created by this criterion for the diagnosis was recognized subsequently, and "erectile or ejaculatory dysfunction" (APA, 1994, p. 449) was therefore added to the final version. However, this addition was not based on any studies with males, the derivation and validation of the *DSM–IV* criteria set again being confined to females (Yutzy et al., 1995). In his critique of the *DSM–III* criteria set, Cloninger (1987) had noted that "the criteria for men were based on the work with women" (p. 254) and indicated that "data are needed to validate the final criteria adopted in *DSM-III* for both men and women" (p. 255). However, the analyses provided for *DSM–IV* somatization disorder continued to be confined to females (Cloninger, 1996; Cloninger & Yutzy, 1993; Yutzy et al., 1995).

Research on males with somatization disorder has been conducted. For example, Cloninger, Martin, et al. (1986) indicated an absence of a family history of somatization disorder in males with the disorder and considerable heterogeneity in the comorbid symptomatology. Cloninger, von Knorring, Sigvardsson, and Martin (1986) referred to the male variant as "asthenic somatization," characterized by less diverse complaints and more emphasis on fatigue and weakness than on pain. Cloninger, Martin, et al. (1986) concluded that "these observations do not provide any support for regarding somatization in men as a homogeneous syndrome or a discrete clinical disorder" (p. 877). These conclusions, however, were disputed subsequently by Golding, Smith, and Kashner (1991), who suggested that males with the disorder either fail to seek treatment or are misdiagnosed as antisocial or malingering. They noted that they had much more success in obtaining cases from newspaper ads than from clinician referrals. They concluded that "clinically, male and female patients with Somatization Disorder showed more similarities than differences" (Golding et al., 1991, p. 234).

It may indeed be the case that "Somatization Disorder occurs only rarely in men" (APA, 1994, p. 447) and that the inclusion of female reproductive symptomatology is necessary for—and increases the validity of—the diagnosis in females; however, these gender-specific criteria may also inhibit, if not prevent, any recognition of a future change in the prevalence rate within males. Research on the epidemiology of the disorder therefore continues to use a criteria set that is biased against making the diagnosis in males, complicating any identification of the disorder in males. The World Health Organization's (1992) diagnostic criteria for this disorder are not as gender specific. Menstrual symptoms are included, but it is also stated that "gastrointestinal sensations (pain, belching, regurgitation, vomiting, nausea, etc.) and abnormal skin sensations (itching, burning, tingling, numbness, soreness, etc.) and blotchiness are among the commonest" symptoms (World Health Organization, 1992, p. 162). Perhaps as a result, the disorder is diagnosed more frequently in males within other cultures (Kirmayer & Taillefer, 1997).

Other disorders. The difficulties in developing unbiased diagnostic criteria for disorders that are expressed differently across males and females or are associated with gender-related

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

behaviors are evident in many other mental disorders, including gender identity disorder, dependent personality disorder, and schizophrenia. For example, separate diagnostic criteria are provided for males and females for the diagnosis of gender identity disorder because the disorder does appear differently in boys and girls: "in boys, assertion that his penis or testes are disgusting or will disappear or assertion that it would be better not to have a penis, or aversion toward rough-and-tumble play and rejection of male stereotypical toys, games, and activities; in girls, rejection of urinating in a sitting position, assertion that she has or will grow a penis, or assertion that she does not want to grow breasts or menstruate, or marked aversion toward normative feminine clothing" (APA, 1994, p. 537). Wanting to grow breasts is not characteristic of gender identity disorder in girls, and marked aversion toward feminine clothes is not particularly characteristic of gender identity disorder in boys. It is unclear whether the provision of separate criteria in this instance avoids a bias (by not using the same differentially valid criteria for both sexes; Robins, 1986, 1991; Zoccolillo, 1993) or is, in fact, biased (by creating a different disorder in males and females that has a different threshold for diagnosis; Wakefield, 1987; Zahn-Waxler, 1993).

Kaplan (1983b) had argued that the "*DSM-III* singles out for scrutiny and therefore diagnosis the ways in which women express dependency but not the ways in which men express dependency" (p. 789). Members of the *DSM–IV* task force agreed that the higher rate of dependent personality disorder in women did reflect in part "masculine biases regarding what is healthy behavior and the failure to recognize masculine forms of dependent behavior" (Frances et al., 1995, p. 377). L. E. A. Walker (1994), for example, suggested that "men who rely on others to maintain their homes and take care of their children are . . . expressing personality-disordered dependency behaviors" (p. 25). Frances et al. (1995) admonished clinicians "not to miss stereotypically masculine forms of dependency expressed through domineering behavior, ordering others to help rather than demanding or pleading" (p. 377). Bornstein (1996) suggested that the inclusion of diagnostic criteria that recognized a denial of dependency would result in a more equivalent sex ratio. Kaplan (1983b) suggested that dependent widowed men seek a new spouse to be taken care of, whereas widowed women seek a new spouse to care for. Phillips and Gunderson (1994), however, suggested that dependent personality disorder has a differential sex prevalence rate because it is related to different social cultural expectations that contribute to the etiology of the disorder. Dependency on others may be encouraged in females more so than in males (Gilligan, 1982; Hamilton & Jensvold, 1992). "Thus, Dependent Personality Disorder may represent an exaggerated and maladaptive variant of normal dependency" (Phillips & Gunderson, 1994, p. 720).

Kraepelin (1919/1971) characterized schizophrenia as a disorder of young men. It is now suggested that the sex ratio for schizophrenia is approximately equal (APA, 1994), although there are important differences between the sexes in the course and presentation of the disorder. For example, "women are more likely to have a later onset, more prominent mood symptoms, and a better prognosis" (APA, 1994, p. 281). The presence of the more prominent mood symptomatology in females with schizophrenia complicates the differential sex prevalence of the associated diagnosis of schizoaffective disorder (Goldstein, 1995). There may not necessarily be a higher absolute ratio of females to males with schizoaffective disorder but, rather, a higher ratio only relative to the ratio of females to males with schizophrenia. "Compared with Schizophrenia, Schizoaffective Disorder probably occurs more often in women" (APA, 1994, p. 294). Goldstein (1995) suggested that females with schizophrenia express more affective symptoms, paranoia, and auditory hallucinations, whereas males with schizophrenia have more negative symptoms, such as flat or constricted affect, anhedonia, and social withdrawal. Therefore, "broader definitions of Schizophrenia with respect to the boundary with Mood Disorders . . . yield a higher female-to-male ratio than the relatively narrow construct of Schizophrenia used in this manual" (APA, 1994, p. 282). Or the failure to recognize the greater prominence of mood symptoms in women with schizophrenia contributes to an overdiagnosis of schizoaffective disorder in women relative to men.

It may appear unbiased to have a criteria set for schizophrenia that results in an equal sex ratio, but an equal sex ratio would reflect a bias if the criteria set fails to recognize actual sex differences (Widiger & Spitzer, 1991). "In general, recent studies have shown that the incidence is higher among men than among women, especially when stringent diagnostic criteria for schizophrenia are applied" (Goldstein, 1995, p. 191). Kraepelin's (1919/1971) characterization of schizophrenia as a disease of young men may be correct. The trend in research on schizophrenia is toward a more narrowly defined criteria set (Frances et al., 1995; Spitzer, Williams, & Skodol, 1980), which will in turn favor the male variant. True schizophrenia might then be more prevalent among males, whereas females would be more likely to have the boundary condition of schizoaffective disorder.

Recommendations

Most of the mental disorders diagnosed within the *DSM–IV* do appear to have significant differential sex prevalence rates (Anthony et al., 1995; Corbitt & Widiger, 1995; Eme & Kavanaugh, 1995; Wolk & Weissman, 1995). It would be unlikely to find a virtually equivalent sex ratio (Meehl, 1967), the differences that do occur will often be substantial, and the rates provided in the *DSM–IV* are likely to be, for the most part, reasonable estimates based on the existing research. However, the most common (and often heated) critiques of the *DSM–IV* have concerned charges of sex bias (Brown, 1992; Caplan, 1991, 1995; Gallant & Hamilton, 1988; Kaplan, 1983a, 1983b; Landrine, 1989; Ross et al., 1995), and these charges may not be resolved until the methodological and conceptual problems discussed earlier are adequately addressed.

Informative data concerning differential sex prevalence rates can be obtained from research within clinical settings. The most parsimonious explanation for a significant sex ratio across different clinical settings is the presence of a differential sex prevalence rate for the disorder. However, researchers do need to be sensitive to the many possible alternative reasons a differential sex prevalence rate can occur within a particular clinical setting. The most informative findings will be provided by epidemiological studies involving representative samples of respective populations (Gaub & Carlson, in press). Probability-based samplings

of community populations are particularly necessary (Sher & Trull, 1996), given the many alternative reasons for differential prevalence rates within clinical settings. Regrettably, however, few such studies have been conducted, and these investigations have been limited with respect to the disorders included. As a result, very few of the statements concerning differential sex prevalence rates in the *DSM–IV* were (or could have been) based on adequate probability-based community samples.

For example, the highly informative NIMH Epidemiological Catchment Area (ECA) study provided data on less than a fourth of the 101 disorders for which information regarding differential sex prevalence is provided in the *DSM–IV*. No data were obtained for any disorders of childhood or for any dissociative, sexual and gender identity, factitious, eating, or impulse-control disorders (data were obtained on some of these disorders within a subset of sites). Data were obtained for only one of the personality disorders (antisocial) and for only one of the somatoform disorders (somatization). In addition, the ECA study was confined to five designated catchment areas that contained at least 200,000 residents. The ECA sampling was not intended to provide national estimates of prevalence rates but to assess, instead, the prevalence within particular settings and the proportion receiving mental health services. "The goals of the ECA study to include institutional respondents and clinical reappraisals made it necessary to carry out the ECA study in a small number of local samples" (Kessler et al., 1994, p. 10). Discussion of these limitations of the ECA study is not intended to be a criticism of its design or implementation; the findings were indeed highly informative. The study provided "the most comprehensive report ever assembled on the prevalence rates of mental disorders in the United States" (Regier & Robins, 1991, p. 1). The limitations in sampling and coverage suggest only the need for additional, expanded versions of this study.

The National Comorbidity Survey (NCS; Kessler et al., 1994) built on the methodology of the ECA study. It used a stratified probability sampling procedure to obtain a nationally representative sample of 15–54-year-olds within the United States. "The NCS is the first survey to administer a structured psychiatric interview to a representative national sample in the United States" (Kessler et al., 1994, p. 8). However, the coverage was even more limited than in the ECA study, being confined to only 14 of the 101 disorders within the *DSM–IV* (see Tables 1–4) for which information regarding differential sex prevalence rates was provided.

The deficit of epidemiologic data for childhood disorders is being addressed in part by the NIMH Methods for the Epidemiology of Child and Adolescent Mental Disorders (MECA) study (Lahey et al., 1996). Youths in the MECA study are 9–17 years old; there is a flat distribution of age and a nearly equal rate of boys and girls (as in the general population). The sample was not selected to be fully representative of youths in the United States. "The MECA sample cannot be said to be representative of youths with mental disorders in the United States" (Goodman et al., 1997, p. 10). However, the sampling was not biased by the presence versus absence of a mental disorder or by the use of a mental health service. One of its explicit aims is to determine whether there are differences between youths who do and do not receive mental health services, and an initial report by Goodman et al. (1997) indicates that, indeed, prepubertal boys with mental disorders are more likely to receive treatment than prepubertal girls. However, Goodman et al. emphasized that they could not determine whether the clinical populations were in fact unrepresentative of the disorder within the population. "Only a representative national sample can answer such a question" (Goodman et al., 1997, p. 10).

If a disorder does occur much more often in one of the two sexes, it remains important to include or represent the minority sex within clinical samples to ensure that the findings are not specific to (biased in favor of) the majority sex (NIH, 1994). It is not unusual in current research to exclude members of the minority sex to obtain more homogenous, prototypic cases. This research remains highly informative, but there is the potential for the findings to be less applicable to the minority sex, particularly for disorders that are associated with gender or are expressed differently across the sexes. In fact, there may be very few mental disorders for which there are no significant differences in how the disorder appears or is expressed across males and females, as suggested in the studies cited earlier concerning conduct disorder, ADHD, antisocial personality disorder, factitious disorder, histrionic personality disorder, dependent personality disorder, somatization disorder, gender identity disorder, substance-related disorders, and schizophrenia.

A failure to consider differences in how a disorder appears or is expressed across the sexes is likely to result in the development of diagnostic criteria that are not equally valid for the two sexes, complicating substantially the obtainment of valid prevalence rates in epidemiologic studies. The preference in the *DSM–IV* is for the development of gender-neutral diagnostic criteria, or diagnostic criteria that are the same, and presumably equally valid, for both sexes. Separate male and female diagnostic criteria are provided in the *DSM–IV* only for the sexual and gender identity disorders (APA, 1994). Nevertheless, the potential inadequacy of using the same criteria for both sexes is becoming increasingly evident for many other disorders.

A variety of approaches to identifying empirically the presence of gender bias have been developed within psychometric research on assessment instruments. This psychometric research has focused primarily on the identification of ethnic, cultural, or gender biases within instruments for the assessment of mental disorders, but the principles and techniques of the research are readily applicable to the diagnostic criteria themselves (Blashfield & Livesley, 1991; Morey, 1991; Nelson-Gray, 1991; Widiger et al., 1984). Diagnostic criteria sets are essentially fallible indicators of the presence of a particular mental disorder (Meehl, 1986; Widiger & Trull, 1991), and one source of fallibility would be the presence of gender biases.

For example, a predominant approach to identifying biases within assessment instruments is determining whether there are differences across demographic groups in the accuracy with which test scores can predict relevant extratest characteristics (e.g., Kline & Lachar, 1992; Timbrook & Graham, 1994). One could likewise assess for male–female differences with respect to the strength of the relationship of diagnostic criteria with external validators (e.g., course, treatment responsivity, or family history). Studies have compared the validity of diagnostic criteria sets for males and females (e.g., Frick et al., 1994; Rutherford et al., 1995), but there are surprisingly few such investigations. Many studies have compared males and females

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

with respect to diagnostic criteria and external validators of a disorder (Gaub & Carlson, in press; Nolen-Hoeksema, 1995), but few have considered whether the diagnostic criteria vary across males and females with respect to the strength of their relationship with these external validators.

Various taxometric methods for both categorical (Meehl, 1995) and dimensional (Livesley, Schroeder, Jackson, & Jang, 1994) diagnostic constructs would be applicable. One may find that the association of a diagnostic criterion to an external validator, the association of the criteria to one another, or the factor structure of a criteria set varies across males and females. Such methods have been used to assess the validity of assessment instruments and theoretical models of pathology across the genders (Nolen-Hoeksema, 1995), but they have not been applied to the question of the validity of the diagnostic criteria sets across the genders. For example, if one discovered virtually no relationship of a criterion to a validator within one sex but a substantial association with the validator within mixed-sex groups, this might suggest that the criterion was successful in diagnosing the disorder primarily through its ability to distinguish between the sexes.

Lindsay and Widiger (1995), for instance, indicated that some items from self-report inventories for the assessment of personality disorders were uncorrelated with other indicators of the disorder when sex of participant was controlled, but these items were correlated with gender. They suggested that these items were indicators of the presence of the disorder primarily because the disorder was itself gender related. Some of these items were perhaps blatantly sex biased, because they were indicators of normal rather than abnormal functioning (e.g., responding false to a statement that one frequently had sex with many persons that did not mean much to oneself, used as an indicator for a dependent personality disorder). It is unlikely that the diagnostic criteria for any of the *DSM–IV* mental disorders would be biased to this degree, but it is possible—and, perhaps, even likely—that diagnostic criteria will vary significantly for males and females in their validity, as indicated, for example, by their correlations with external validators within each sex. Such findings might not necessarily suggest a sex bias, because they could reflect different structural associations of the disorder for males and females (Drasgow & Kanfer, 1985; Ozer & Reise, 1994). However, items that vary across males and females in their ability to diagnose the presence of a disorder would involve more diagnostic errors (either false positive or false negative) for one sex than for the other, and in this sense they could be said to be gender biased (Widiger & Spitzer, 1991). The diagnosis would not be entirely or always invalid for that sex, but it would be sufficiently less valid such that some form of correction to the criteria set should be made, such as deleting the problematic diagnostic criteria (Bornstein, 1997; Kaplan, 1983b; Wakefield, 1987; Widiger & Spitzer, 1991), lowering the threshold for one sex (Barkley, 1995a; Cloninger, Martin, et al., 1986; Rutherford et al., 1995), or providing alternative criteria sets (Masters & Johnson, 1982; Zoccolillo, 1993).

Item response theory analyses would be particularly informative in identifying potential gender biases within diagnostic criteria sets. Such analyses are traditionally used to assess whether items within an assessment instrument vary in their discriminative validity at different levels of the construct being assessed (Hambleton, Swaminathan, & Rogers, 1991). However, they can also be used to assess whether the discriminative validity of an item varies for males and females at the same level of the construct being assessed (Embretson, 1996). For example, Santor, Ramsay, and Zuroff (1994) assessed male–female differences on items from the Beck Depression Inventory (Beck, Ward, Mendelson, Mock, & Erbaugh, 1961). Women are diagnosed with depression much more frequently than men (APA, 1994), but finding an overall mean difference between men and women could reflect a true mean difference rather than a bias (Nolen-Hoeksema, 1995). "Item bias can be detected only when individuals have been equated along some continuum such that responses to items can be examined for individuals who are equally depressed" (Santor et al., 1994, p. 256). Santor et al. used item response theory analyses to examine whether men and women at equal levels of depression responded differentially to individual items from the Beck Depression Inventory. There were very few such differences, suggesting that the higher level of depression in women on this instrument was due largely to actual differences in their severity of depression. However, one notable exception occurred for the item concerned with body image dissatisfaction. Women at the same level of depression as men were more likely to endorse this item, suggesting that its endorsement by women was due in part to factors associated with their gender, irrespective of their level of depression. Women's endorsement of this item was also still due in part to their level of depression, but "scores from this item would overestimate the degree of depression in women with respect to men" (Santor et al., 1994, p. 261).

Comparable analyses of the individual items from the *DSM–IV* diagnostic criteria sets would be equally informative. The findings for some diagnostic criteria might not be surprising (e.g., a history of at least one sexual or reproductive symptom other than pain for the diagnosis of somatization disorder or forcing someone into a sexual activity for the diagnosis of conduct disorder), but we anticipate that many unexpected and equally instructive findings could be obtained for many of the diagnostic criteria sets within the *DSM–IV*. Such item response theory analyses will not be without their limitations (Embretson, 1996). For example, the anchoring items with which males and females are equated with respect to the disorder being assessed might themselves be systematically biased in a manner comparable to the item being analyzed. However, an appreciation of the complexity and ambiguity of item response theory analyses of *DSM–IV* diagnostic criteria sets will represent an even further increase in understanding of the relationship of gender to diagnosis.

A problematic issue for the authors of the fifth edition of the *DSM* (*DSM–V*) will be whether to revise existing diagnostic criteria to improve their gender neutrality, develop alternative criteria sets for males and females, or provide alternative diagnostic thresholds. Individual items are likely to perform quite differently in males and females for disorders that are related to gender (e.g., histrionic personality disorder) or are expressed differently in males and females (e.g., somatization disorder). Gender-neutral criteria may be difficult to achieve for many of these disorders in the *DSM–V*. An alternative approach would be the provision of separate criteria for males and females, comparable to the approach used already for the sexual and

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

gender identity disorders (i.e., rather than use of the same criteria with different validity for males and females, use of different criteria with the same, or at least more comparable, validity for males and females).

Zoccolillo and Rogers (1991), for example, reported a course for conduct disorder in girls that was comparable to the course obtained with boys when the disorder was diagnosed in girls with such symptoms as rule violations at home and school, substance abuse, prostitution, chronic lying, running away from home overnight, and poor school performance. These girls would not have met the *DSM–III–R* (or *DSM–IV*) criteria for conduct disorder, because they lacked the symptoms of vandalism, fire setting, burglary, use of a weapon in fights, stealing with confrontation of a victim, and rape. However, criteria that are substantially different across sexes may no longer be diagnosing the same disorder (Zahn-Waxler, 1993). If a disorder has a known specific etiology and pathology, the presence of a common disorder with different symptomatology across sexes could be established through observation of the interaction of gender-related variables with the etiology. In the absence of a known specific etiology, one can also use taxometric techniques to verify the presence of a common—or at least sufficiently comparable—latent class taxon within each sex (Meehl, 1995). However, mental disorders will also have complex, multifactorial etiologies, contributing to substantial heterogeneity in symptomatology even for members of the same sex. The point at which the interaction with gender-related biogenetic and environmental variables warrants the recognition of a distinct disorder for the two sexes will, at times, be unclear. The single *DSM–IV* diagnosis of ADHD, with the subtypes of predominantly inattentive, predominantly hyperactive–impulsive, and combined (APA, 1994), may instead represent two different disorders (Barkley, 1995b; Lahey, Carlson, & Frick, 1997; Paternite, Loney, & Roberts, 1996), with the sex differences confined largely to the hyperactive variant(s). Conduct disorder in girls may indeed be so different from conduct disorder in boys in etiology, pathology, phenomenology, course, or treatment (Eme & Kavanaugh, 1995; Goodman & Kohlsdorf, 1994; Robins, 1986, 1991; Zoccolillo, 1993) that it might be useful to consider them as different conditions, comparable to the many differences in orgasmic disorder across men and women (McConaghy, 1997). The same issues apply in the decision of whether to develop culturally neutral diagnostic criteria, alternative diagnostic criteria for the same disorders across cultures, or culture-specific (culture-bound) syndromes (Culbertson, 1997; Mezzich, Kleinman, Fabrega, & Parron, 1996).

Zahn-Waxler's (1993) concern that different criteria will lower the threshold for the diagnosis of conduct disorder in girls might be addressed by requiring that the alternative criteria sets at least be equated with respect to their level of dysfunction associated with the threshold for the diagnosis. Psychometrically, this is comparable to ensuring not only an equivalence with respect to the strength of the relationship of the diagnostic indicators with external validators (i.e., equivalent slopes with respect to predictive validity) but also an equivalence with respect to the cutoff points used for assessment decisions (i.e., equivalence with respect to the intercept).

The diagnostic thresholds for many of the disorders in the *DSM–IV* do appear to be somewhat arbitary (Clark, Watson, & Reynolds, 1995). For example, the requirements of 3 or more of the 15 criteria for conduct disorder, 5 or more of the 9 criteria for borderline personality disorder, 5 or more of the 11 criteria for premenstrual dysphoric disorder, and 2 or more of the 5 criteria for dysthymia are, for the most part, meaningful and well reasoned; however, they are not based on a qualitative demarcation between the presence and absence of a known pathology (Widiger, 1997). The absence of a nonarbitrary point of demarcation complicates substantially the resolution of controversies concerning gender biases, because there is no absolute standard for a mental disorder diagnosis to identify conclusively whether a threshold is being placed too low or too high for females or for males (Faraone & Tsuang, 1994; Spitzer, 1983; Widiger & Spitzer, 1991). In the absence of such a standard, it would be useful to know whether the thresholds are relatively higher or lower across different mental disorders, particularly those disorders for which there are notable differences in sex prevalence rates. Some disorders will, of course, be associated with more dysfunction than others (e.g., schizophrenia vs. hypoactive sexual desire disorder), but it may also be arbitrarily easier to receive some diagnoses than others (e.g., less dysfunction is required for the diagnosis). Funtowicz and Widiger (1995) compared male-typed and female-typed personality disorders across 30 indicators of social dysfunction, occupational dysfunction, and personal distress. They reported that less dysfunction was required for some of the male-typed disorders (e.g., obsessive–compulsive and antisocial) than for the female-typed disorders (e.g., dependent and borderline), contrary to the hypothesis that the threshold for the diagnosis of disorders concerning stereotypic feminine traits was lower than the threshold for the diagnosis of stereotypic masculine traits. This research will be particularly informative when there are male and female variants of or subtypes for a particular mental disorder (e.g., gender identity disorder) or when alternative criteria sets are being suggested (e.g., conduct disorder) to assess whether the cutoff points between normality and abnormality are being set at a comparable level for males and females.

In sum, researchers might need to take more to heart the admonition of the National Institute on Alcohol Abuse and Alcoholism and NIH to provide a better representation of the sexes within clinical research. Gender differences do appear to be pervasive across the mental disorders. There may not be a mental disorder for which there are not important gender differences in the manner in which the disorder is expressed. Recognizing and defusing potential gender biases can be exceedingly complex and controversial (Eagly, 1995). However, the commonality of the issues across the full range of mental disorder diagnoses covered in the *DSM–IV* is noteworthy. We suggest that, in the development of future editions of the *DSM*, there be an advisory committee on gender issues whose responsibility it would be to identify, coordinate, and advise the task force on the breadth and depth of gender issues across the full range of mental disorder diagnoses included in and excluded from the *DSM–IV* (e.g., disorders of dyscontrolled anger and aggression within adulthood). The development of the *DSM–IV* included an Advisory Committee on Cross-Cultural Issues (Mezzich et al., 1996). The issues and proposals raised in this article for gender differences are applicable to cultural differences and, perhaps, many other demographic variables, and the advisory committee did

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

indeed consider the possibility of culture-specific diagnostic criteria for some mental disorders. Members of this committee were the principal authors of the text discussion of specific cultural features included for each of the diagnoses in the *DSM–IV*. In contrast, the material on gender was prepared by a respective member of the disorder's work group (typically the same person who prepared the literature review on proposed revisions to the respective diagnostic criteria set published in the *DSM–IV Sourcebook;* i.e., Widiger, Frances, Pincus, Ross, & First, 1997). The decisions that were made with respect to the text description of each disorder were, for the most part, based on thorough and conscientious considerations of the empirical research; however, the data necessary for providing the most informed and accurate descriptions of the gender nature of each disorder may be lacking.

Consensus and uniformity would, of course, be difficult to obtain with respect to issues concerning sex bias (Eagly, 1995). Viewpoints can be very strong and, at times, even volatile (Caplan, 1995; Eagly, 1995; Frances et al., 1995; Ross et al., 1995). Joint conferences or a task force on sex bias in clinical diagnosis, with representatives from different disorders and perspectives, might be useful in stimulating discussion and collaboration across domains of research and across the different perspectives and positions on this controversy. The NIMH funded the *DSM–IV* Advisory Committee on Cross-Cultural Issues (Mezzich et al., 1996). Such a conference or committee would be helpful in fostering greater communication and collaboration and in guiding the development of more informative research before publication of the next edition of the *DSM*. In any case, we hope that this article will itself be fruitful in stimulating additional appreciation and communication across the full range of psychopathology of the issues concerning sex bias and differential sex prevalence.

References

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (rev. 3rd ed.). Washington, DC: Author.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Anthony, J. C., Arria, A. M., & Johnson, E. O. (1995). Epidemiological and public health issues for tobacco, alcohol, and other drugs. In J. M. Oldham & M. B. Riba (Eds.), *Review of psychiatry* (Vol. 14, pp. 15–49). Washington, DC: American Psychiatric Press.

Barkley, R. A. (1990). *Attention deficit hyperactivity disorder: A handbook for diagnosis and treatment.* New York: Guilford Press.

Barkley, R. A. (1995a). Sex differences in ADHD. *ADHD Report, 3*(1), 1–4.

Barkley, R. A. (1995b). *Taking charge of ADHD: The complete, authoritative guide for parents.* New York: Guilford Press.

Barkley, R. A., Fisher, M., Edelbrock, C. S., & Smallish, L. (1990). The adolescent outcome of hyperactive children diagnosed by research criteria. I: An 8-year prospective follow-up study. *Journal of the American Academy of Child and Adolescent Psychiatry, 29,* 546–557.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 12,* 57–62.

Blanchard, D. C., Griebel, G., & Blanchard, R. J. (1995). Gender bias in the preclinical psychopharmacology of anxiety: Male models for (predominantly) female disorders. *Journal of Psychopharmacology, 9,* 79–82.

Blashfield, R. K., & Livesley, W. J. (1991). Metaphorical analysis of psychiatric classification as a psychological test. *Journal of Abnormal Psychology, 100,* 262–270.

Blume, S. B. (1990). Chemical dependency in women: Important issues. *American Journal of Drug and Alcohol Abuse, 16,* 297–307.

Bornstein, R. F. (1996). Sex differences in dependent personality disorder prevalence rates. *Clinical Psychology: Science and Practice, 3,* 1–12.

Bornstein, R. F. (1997). Dependent personality disorder in the DSM–IV and beyond. *Clinical Psychology: Science and Practice, 4,* 175–187.

Breslau, N., Davis, G. C., Andreski, P., & Peterson, E. (1991). Traumatic events and posttraumatic stress disorder in an urban population of young adults. *Archives of General Psychiatry, 48,* 216–222.

Brett, P. J., Graham, K., & Smythe, C. (1995). An analysis of specialty journals on alcohol, drugs, and addictive behaviors for sex bias in research methods and reporting. *Journal of Studies on Alcohol, 56,* 24–34.

Brown, L. S. (1992). A feminist critique of the personality disorders. In L. S. Brown & M. Ballou (Eds.), *Personality and psychopathology. Feminist reappraisals* (pp. 206–228). New York: Guilford Press.

Buss, D., & Schmitt, D. (1993). Sexual strategies theory: An evolutionary perspective on human mating. *Psychological Review, 100,* 204–232.

Caplan, P. J. (1991). How do they decide who is normal? The bizarre, but true, tale of the DSM process. *Canadian Psychology, 32,* 162–170.

Caplan, P. J. (1995). *They say you're crazy. How the world's most powerful psychiatrists decide who's normal.* Reading, MA: Addison-Wesley.

Chodoff, P. (1982). Hysteria and women. *American Journal of Psychiatry, 139,* 545–551.

Clark, L. A., Watson, D., & Reynolds, S. (1995). Diagnosis and classification of psychopathology: Challenges to the current system and future directions. *Annual Review of Psychology, 46,* 121–153.

Cloninger, C. R. (1987). Diagnosis of somatoform disorders: A critique of DSM–III. In G. L. Tischler (Ed.), *Diagnosis and classification in psychiatry: A critical appraisal of DSM–III* (pp. 243–259). New York: Cambridge University Press.

Cloninger, C. R. (1996). Somatization disorder. In T. A. Widiger, A. J. Frances, H. A. Pincus, R. Ross, M. B. First, & W. W. Davis (Eds.), *DSM–IV sourcebook* (Vol. 2, pp. 885–892). Washington, DC: American Psychiatric Association.

Cloninger, C. R., Martin, R. L., Guze, S. B., & Clayton, P. J. (1986). A prospective follow-up and family study of somatization in men and women. *American Journal of Psychiatry, 143,* 873–878.

Cloninger, C. R., von Knorring, A., Sigvardsson, S., & Martin, R. L. (1986). Symptom patterns and causes of somatization in men: Part II, genetic and environmental independence from somatization in women. *General Epidemiology, 3,* 171–185.

Cloninger, C. R., & Yutzy, S. (1993). Somatoform and dissociative disorders: A summary of changes for DSM–IV. In D. L. Dunner (Ed.), *Current psychiatric therapy* (pp. 310–313). Philadelphia: W. B. Saunders.

Cooper, A. M. (1987). Histrionic, narcissistic, and compulsive personality disorders. In G. L. Tischler (Ed.), *Diagnosis and classification in psychiatry: A critical appraisal of DSM–III* (pp. 290–299). New York: Cambridge University Press.

Corbitt, E. M., & Widiger, T. A. (1995). Sex differences among the personality disorders: An exploration of the data. *Clinical Psychology: Science and Practice, 2,* 225–238.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Culbertson, F. M. (1997). Depression and gender. An international review. *American Psychologist, 52,* 25–31.

Davidson, J. R. T., & March, J. S. (1997). Traumatic stress disorders. In A. Tasman, J. Kay, & J. A. Lieberman (Eds.), *Psychiatry* (Vol. 2, pp. 1085–1099). Philadelphia: W. B. Saunders.

Drasgow, F., & Kanfer, R. (1985). Equivalence of psychological measurement in heterogeneous populations. *Journal of Applied Psychology, 70,* 662–680.

Eagly, A. H. (1995). The science and politics of comparing women and men. *American Psychologist, 50,* 145–158.

Embretson, S. E. (1996). The new rules of measurement. *Psychological Assessment, 8,* 341–349.

Erne, R. F. (1992). Selective female affliction in the developmental disorders: A literature review. *Journal of Clinical Child Psychology, 21,* 354–364.

Erne, R. F., & Kavanaugh, L. (1995). Sex differences in conduct disorder. *Journal of Clinical Child Psychology, 24,* 406–426.

Emrick, C. D. (1974). A review of psychologically oriented treatment of alcoholism. I. The use and interrelationships of outcome criteria and drinking behavior following treatment. *Quarterly Journal of Studies on Alcohol, 35,* 523–549.

Faraone, S. V., & Tsuang, M. T. (1994). Measuring diagnostic accuracy in the absence of a "gold standard." *American Journal of Psychiatry, 151,* 650–657.

Feingold, A. (1994). Gender differences in personality: A meta-analysis. *Psychological Bulletin, 116,* 429–456.

Frances, A. J., First, M. B., & Pincus, H. A. (1995). *DSM–IV guidebook.* Washington, DC: American Psychiatric Press.

Frick, P. J., Lahey, B. B., Applegate, B., Kerdyk, L., Ollendick, T., Hynd, G. W., Garfinkel, B., Greenhill, L., Biederman, J., Barkley, R. A., McBurnett, K., Newcorn, J., & Waldman, L. (1994). DSM–IV field trials for the disruptive behavior disorders: Symptom utility estimates. *Journal of the American Academy of Child and Adolescent Psychiatry, 33,* 529–539.

Funtowicz, M. N., & Widiger, T. A. (1995). Sex bias in the diagnosis of personality disorders: A different approach. *Journal of Psychopathology and Behavioral Assessment, 17,* 145–165.

Gallant, S. J., & Hamilton, J. A. (1988). On a premenstrual psychiatric diagnosis: What's in a name? *Professional Psychology: Research and Practice, 19,* 271–278.

Gannon, L., Luchetta, T., Rhodes, K., Pardie, L., & Segrist, D. (1992). Sex bias in psychological research: Progress or complacency? *American Psychologist, 47,* 389–396.

Garb, H. N. (1997). Race bias, social class bias, and gender bias in clinical judgment. *Clinical Psychology: Science and Practice, 4,* 99–120.

Gaub, M., & Carlson, C. L. (in press). Gender differences in ADHD: A meta-analysis and critical review. *Journal of the American Academy of Child and Adolescent Psychiatry.*

Gilligan, C. (1982). *In a different voice: Psychological theory and women's development.* Cambridge, MA: Harvard University Press.

Golding, J. M., Smith, R., & Kashner, T. M. (1991). Does somatization disorder occur in men? *American Journal of Psychiatry, 48,* 231–235.

Goldstein, J. M. (1993). Sampling biases in studies of gender and schizophrenia: A reply. *Schizophrenia Bulletin, 19,* 9–14.

Goldstein, J. M. (1995). The impact of gender on understanding the epidemiology of schizophrenia. In M. V. Seeman (Ed.), *Gender and psychopathology* (pp. 159–199). Washington, DC: American Psychiatric Press.

Goldstein, J. M., & Tsuang, M. T. (1990). Gender and schizophrenia: An introduction and synthesis of findings. *Schizophrenia Bulletin, 16,* 179–183.

Good, G. E., & Wood, P. K. (1995). Male gender role conflict, depression, and help seeking: Do college men face double jeapardy? *Journal of Counseling and Development, 74,* 70–75.

Goodman, S. H., & Kohlsdorf, B. (1994). The developmental psychopathology of conduct problems: Gender issues. In D. C. Fowles, P. Sutker, & S. H. Goodman (Eds.), *Progress in experimental personality and psychopathology research* (Vol. 15, pp. 121–161). New York: Springer.

Goodman, S. H., Lahey, B. B., Fielding, B., Dulcan, M., Narrow, W., & Regier, D. (1997). Representativeness of clinical samples of youths with mental disorders: A preliminary population-based study. *Journal of Abnormal Psychology, 106,* 3–14.

Guze, S. B. (1967). The diagnosis of hysteria: What are we trying to do? *American Journal of Psychiatry, 124,* 491–498.

Guze, S. B., Woodruff, R. A., & Clayton, P. J. (1972). Sex, age, and the diagnosis of hysteria (Briquet's syndrome). *American Journal of Psychiatry, 129,* 745–748.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory.* Newbury Park, CA: Sage.

Hamilton, J. A., & Jensvold, M. (1992). Personality, psychopathology, and depressions in women. In L. S. Brown & M. Ballou (Eds.), *Personality and psychopathology. Feminist reappraisals* (pp. 116–143). New York: Guilford Press.

Helzer, J. E., Robins, L. N., & McEvoy, L. (1987). Post-traumatic stress disorder in the general population. *New England Journal of Medicine, 317,* 1630–1634.

Hirschfeld, R. M. A., Shea, M. T., & Weise, R. (1991). Dependent personality disorder: Perspectives for DSM–IV. *Journal of Personality Disorders, 5,* 135–149.

Hughes, T. L. (1990). Evaluating research on chemical dependency among women: A women's health perspective. *Family and Community Health, 13,* 35–46.

Huselid, R. F., & Cooper, M. L. (1994). Gender roles as mediators of sex differences in expressions of pathology. *Journal of Abnormal Psychology, 103,* 595–603.

Kaminsky, M. J., & Slavney, P. R. (1976). Methodology and personality in Briquet's syndrome: A reappraisal. *American Journal of Psychiatry, 133,* 85–88.

Kaplan, M. (1983a). The issue of sex bias in DSM–III. Comments on the articles by Spitzer, Williams, and Kass. *American Psychologist, 38,* 802–803.

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

Keenan, K., & Shaw, D. (1997). Developmental and social influences on young girls' early problem behavior. *Psychological Bulletin, 121,* 95–113.

Kessler, R. C., McGonagle, K. A., Zhao, S., Nelson, C. B., Hughes, M., Eshleman, S., Wittchen, H., & Kendler, K. (1994). Lifetime and 12-month prevalence of DSM–III–R psychiatric disorders in the United States. Results from the National Comorbidity Survey. *Archives of General Psychiatry, 51,* 8–19.

Kirmayer, L. J., & Taillefer, S. (1997). Somatoform disorders. In S. M. Turner & M. Hersen (Eds.), *Adult psychopathology and diagnosis* (3rd ed., pp. 333–383). New York:

Kline, R. B., & Lachar, D. (1992). Evaluation of age, sex, and race bias in personality inventory for children (PIC). *Psychological Assessment, 4,* 333–339.

Kraepelin, E. (1971). *Dementia praecox and paraphrenia* (R. M. Barclay, Trans.). New York: Robert E. Krieger. (Original work published 1919).

Lahey, B. B., Applegate, B., McBurnett, K., Biederman, J., Greenhill, L., Hynd, G. W., Barkley, R. A., Newcorn, J., Jensen, P., Richters, J., Garfinkel, B., Kerdyk, L., Frick, P. J., Ollendick, T., Perez, D., Hart, E. L., Waldman, I., & Shaffer, D. (1994). DSM–IV field trials for

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

attention deficit hyperactivity disorder in children and adolescents. *American Journal of Psychiatry, 151,* 1673–1685.

Lahey, B. B., Carlson, C. L., & Frick, P. J. (1997). Attention-deficit disorder without hyperactivity. In T. A. Widiger, A. J. Frances, H. A. Pincus, R. Ross, M. B. First, & W. Davis (Eds.), *DSM–IV sourcebook* (Vol. 3, pp. 163–188). Washington, DC: American Psychiatric Association.

Lahey, B. B., Flagg, E. W., Bird, H. R., Schwab-Stone, M., Canino, G., Dulcan, M. K., Leaf, P. J., Davies, M., Brogan, D., Bourdon, K., Horwitz, S. M., Rubio-Stipec, M., Freeman, D. H., Lichtman, J., Shaffer, D., Goodman, S. H., Narrow, W. E., Weissman, M. M., Kandel, D. B., Jensen, P. S., Richters, J. E., & Regier, D. A. (1996). The NIMH Methods for the Epidemiology of Child and Adolescent Mental Disorders (MECA) study: Background and methodology. *Journal of the American Academy of Child and Adolescent Psychiatry, 35,* 855–864.

Landrine, H. (1989). The politics of personality disorder. *Psychology of Women Quarterly, 13,* 325–339.

Leibenluft, E. (1996). Women with bipolar illness: Clinical and research issues. *American Journal of Psychiatry, 153,* 163–173.

Liederman, J., & Flannery, K. A. (1995). The sex ratios of families with a neurodevelopmentally disordered child. *Journal of Child Psychology and Psychiatry, 36,* 511–517.

Lindsay, K., & Widiger, T. A. (1995). Sex and gender bias in self-report personality disorder inventories: Item analyses of the MCMI–II, MMPI, and PDQ-R. *Journal of Personality Assessment, 65,* 1–20.

Livesley, W. J., Schroeder, M. L., Jackson, D. N., & Jang, K. L. (1994). Categorical distinctions in the study of personality disorder: Implications for classification. *Journal of Abnormal Psychology, 103,* 6–17.

Loeber, R., & Keenan, K. (1994). Interaction between conduct disorder and its comorbid conditions: Effects of age and gender. *Clinical Psychology Review, 14,* 497–523.

Masters, W. H., & Johnson, V. E. (1982). *Human sexuality.* Boston: Little, Brown.

McConaghy, N. (1997). Sexual and gender identity disorders. In S. M. Turner & M. Hersen (Eds.), *Adult psychopathology and diagnosis* (3rd ed., pp. 409–464). New York: Wiley.

Meehl, P. E. (1967). Theory-testing in psychology and physics: A methodological paradox. *Philosophy of Science, 34,* 103–115.

Meehl, P. E. (1986). Diagnostic taxa as open concepts: Meta theoretical and statistical questions about reliability and construct validity in the grand strategy of nosological revision. In T. Millon & G. Klerman (Eds.), *Contemporary directions in psychopathology* (pp. 215–231). New York: Guilford Press.

Meehl, P. E. (1995). Bootstraps taxometrics: Solving the classification problem in psychopathology. *American Psychologist, 50,* 266–275.

Mezzich, J. E., Kleinman, A., Fabrega, H., & Parron, D. L. (Eds.). (1996). *Culture and psychiatric diagnosis: A DSM–IV perspective.* Washington, DC: American Psychiatric Press.

Moffitt, T. E., Caspi, A., Dickson, N., Silva, P., & Stanton, W. (1996). Childhood-onset versus adolescent-onset antisocial conduct problems in males: Natural history from ages 3 to 18 years. *Development and Psychopathology, 8,* 399–424.

Morey, L. C. (1991). Classification of mental disorder as a collection of hypothetical constructs. *Journal of Abnormal Psychology, 100,* 289–293.

Nadeau, K. G. (1995). *A comprehensive guide to attention deficit disorder in adults: Research, diagnosis, and treatment.* New York: Brunner/Mazel.

Nangle, D. W., & Foster, S. (1992). The effects of a positive behavioral context on the social impact of aggressive behavior. *Journal of Abnormal Child Psychology, 20,* 543–553.

National Institutes of Health. (1994). NIH guidelines on the inclusion of women and minorities as subjects in clinical research. *NIH Guide, 23*(10), 1–34.

Nelson-Gray, R. O. (1991). DSM–IV: Empirical guidelines from psychometrics. *Journal of Abnormal Psychology, 100,* 308–315.

Nolen-Hoeksema, S. (1995). Epidemiology and theories of gender differences in unipolar depression. In M. V. Seeman (Ed.), *Gender and psychopathology* (pp. 89–112). Washington, DC: American Psychiatric Press.

Othmer, E., & DeSouza, C. (1985). A screening test for somatization disorder. *American Journal of Psychiatry, 142,* 1146–1149.

Ozer, D. J., & Reise, S. P. (1994). Personality assessment. *Annual Review of Psychology, 45,* 357–388.

Paternite, C., Loney, J., & Roberts, M. (1996). A preliminary validation of subtypes of DSM–IV attention deficit hyperactivity disorder. *Journal of Attention Disorders, 1,* 70–86.

Pfohl, B. (1991). Histrionic personality disorder: A review of available data and recommendations for DSM–IV. *Journal of Personality Disorders, 5,* 150–166.

Phillips, K. A., & Gunderson, J. G. (1994). Personality disorders. In R. E. Hales, S. C. Yudofsky, & J. A. Talbott (Eds.), *Textbook of Psychiatry* (2nd ed., pp. 701–728). Washington, DC: American Psychiatric Press.

Plewes, J. M., & Fagan, J. G. (1994). Factitious disorders and malingering. In R. E. Hales, S. C. Yudofsky, & J. A. Talbott (Eds.), *Textbook of psychiatry* (2nd ed., pp. 623–632). Washington, DC: American Psychiatric Press.

Popper, C. W., & Steingard, R. J. (1994). Disorders usually first diagnosed in infancy, childhood, or adolescence. In R. E. Hales, S. C. Yudofsky, & J. A. Talbott (Eds.), *Textbook of psychiatry* (2nd ed., pp. 729–831). Washington, DC: American Psychiatric Press.

Potts, M. K., Burnam, M. A., & Wells, K. B. (1991). Gender differences in depression detection: A comparison of clinician diagnosis and standardized assessment. *Psychological Assessment, 3,* 609–615.

Regier, D. A., & Robins, L. N. (1991). Introduction. In L. N. Robins & D. A. Regier (Eds.), *Psychiatric disorders in America. The Epidemiologic Catchment Area study* (pp. 1–10). New York: Free Press.

Reich, J. H. (1987). Sex distribution of DSM–III personality disorders in psychiatric outpatients. *American Journal of Psychiatry, 144,* 485–488.

Robins, L. N. (1986). The consequences of conduct disorder in girls. In D. Olweus, M. R. Yarrow, & J. Block (Eds.), *The development of antisocial and prosocial behavior* (pp. 385–414). New York: Academic Press.

Robins, L. N. (1991). Conduct disorder. *Journal of Child Psychology and Psychiatry, 32,* 193–212.

Robins, L. N., & Price, R. K. (1991). Adult disorders predicted by childhood conduct disorder problems: Results from the NIMH Epidemiologic Catchment Area project. *Psychiatry, 54,* 116–132.

Ross, R., Frances, A. J., & Widiger, T. A. (1995). Gender issues in DSM–IV. In J. M. Oldham & M. B. Riba (Eds.), *Review of psychiatry* (Vol. 14, pp. 205–226). Washington, DC: American Psychiatric Press.

Rutherford, M. J., Alterman, A. I., Cacciola, J. S., & Snider, E. C. (1995). Gender differences in diagnosing antisocial personality disorder in methadone patients. *American Journal of Psychiatry, 152,* 1309–1316.

Santor, D. A., Ramsay, J. O., & Zuroff, D. C. (1994). Nonparametric item analyses of the Beck Depression Inventory: Evaluating gender item bias and response option weights. *Psychological Assessment, 6,* 255–270.

Schmidt, C., Klee, L., & Ames, G. (1990). Review and analysis of literature on indicators of women's drinking problems. *British Journal of Addictions, 85,* 179–192.

Schover, L. R. (1996). Female orgasm: Proposals regarding name, criteria, and subtyping. In T. A. Widiger, A. J. Frances, H. A. Pincus, R.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Ross, M. B. First, & W. W. Davis (Eds.), *DSM–IV sourcebook* (Vol. 2., pp. 1115–1128). Washington, DC: American Psychiatric Association.

Shaffer, D. (1994). Attention deficit hyperactivity disorder in adults. *American Journal of Psychiatry, 151,* 633–638.

Shaywitz, S. E., Shaywitz, B. A., Fletcher, J. M., & Escobar, M. D. (1990). Prevalence of reading disability in boys and girls. Results of the Connecticut Longitudinal Study. *Journal of the American Medical Association, 264,* 998–1002.

Shaywitz, S. E., Shaywitz, B. A., Pugh, K. R., Constable, R. T., Skudlarski, P., Fulbright, R. K., Bronen, R. A., Fletcher, J. M., Shankweller, D. P., Katz, L., & Gore, J. C. (1995). Sex differences in the functional organization of the brain for language. *Nature, 373,* 607–609.

Sher, K. J., & Trull, T. J. (1996). Methodological issues in psychopathology research. *Annual Review of Psychology, 47,* 371–400.

Spiegel, D. (1997). Dissociative disorders. In A. Tasman, J. Kay, & J. A. Lieberman (Eds.), *Psychiatry* (Vol. 2, pp. 1156–1172). Philadelphia. W. B. Saunders.

Spitzer, R. L. (1983). Psychiatric diagnosis: Are clinicians still necessary? *Comprehensive Psychiatry, 24,* 399–411.

Spitzer, R. L., Williams, J. B. W., & Skodol, A. E. (1980). DSM–III: The major achievements and an overview. *American Journal of Psychiatry, 137,* 151–164.

Sprock, J., Blashfield, R. K., & Smith, B. (1990). Gender weighting of DSM–III–R personality disorder criteria. *American Journal of Psychiatry, 147,* 586–590.

Task Force on DSM–IV. (1991). *DSM–IV options book: Work in progress.* Washington, DC: American Psychiatric Association.

Timbrook, R. E., & Graham, J. R. (1994). Ethnic differences on the MMPI-2? *Psychological Assessment, 6,* 212–217.

Toneatto, A., Sobell, L. C., & Sobell, M. B. (1992). Gender issues in the treatment of abusers of alcohol, nicotine, and other drugs. *Journal of Substance Abuse, 4,* 209–218.

Vannicelli, M., & Nash, L. (1984). Effect of sex bias on women's studies on alcoholism. *Alcoholism: Clinical and Experimental Research, 8,* 334–336.

Wahl, O. F. (1977). Sex bias in schizophrenia research: A short report. *Journal of Abnormal Psychology, 86,* 195–198.

Wahl, O. F., & Hunter, J. (1992). Are gender effects being neglected in schizophrenia research? *Schizophrenia Bulletin, 18,* 313–317.

Wakefield, J. C. (1987). Sex bias in the diagnosis of primary orgasmic dysfunction. *American Psychologist, 42,* 464–471.

Walker, E. F., & Lewine, R. J. (1993). Sampling biases in studies of gender and schizophrenia. *Schizophrenia Bulletin, 19,* 1–7.

Walker, L. E. A. (1994). Are personality disorders gender biased? In S. A. Kirk & S. D. Einbinder (Eds.), *Controversial issues in mental health* (pp. 22–29). New York: Allyn & Bacon.

Webster-Stratton, C. (1996). Early-onset conduct problems: Does gender make a difference? *Journal of Consulting and Clinical Psychology, 64,* 540–551.

Wender, P. H. (1995). *Attention-deficit hyperactivity disorder in adults.* New York: Oxford University Press.

Widiger, T. A. (1995). Deletion of self-defeating and sadistic personality disorders. In W. J. Livesley (Ed.), *The DSM–IV personality disorders* (pp. 359–373). New York: Guilford Press.

Widiger, T. A. (1997). Mental disorders as discrete clinical conditions: Dimensional versus categorical classification. In S. M. Turner & M. Hersen (Eds.), *Adult psychopathology and diagnosis* (3rd ed., pp. 3–23). New York: Wiley.

Widiger, T. A., Frances, A. J., Pincus, H. A., Ross, R., & First, M. B. (Eds.). (1997). *DSM–IV sourcebook* (Vol. 3). Washington, DC: American Psychiatric Association.

Widiger, T. A., Hurt, S., Frances, A. J., Clarkin, J., & Gilmore, M. (1984). Diagnostic efficiency and DSM–III. *Archives of General Psychiatry, 41,* 1005–1012.

Widiger, T. A., Mangine, S., Corbitt, E. M., Ellis, C. G., & Thomas, G. V. (1995). *Personality disorder interview-IV. A semistructured interview for the assessment of personality disorders.* Odessa, FL: Psychological Assessment Resources.

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

Widiger, T. A., & Trull, T. J. (1991). Diagnosis and clinical assessment. *Annual Review of Psychology, 42,* 109–133.

Wilke, D. (1994). Women and alcoholism: How a male-as-norm bias affects research, assessment, and treatment. *Health and Social Work, 19,* 29–35.

Wolk, S. I., & Weissman, M. M. (1995). Women and depression: An update. In J. M. Oldham & M. B. Riba (Eds.), *Review of psychiatry* (Vol. 14, pp. 227–259). Washington, DC: American Psychiatric Press.

Wool, C. A., & Barsky, A. J. (1994). Do women somatize more than men? Gender differences in somatization. *Psychosomatics, 35,* 445–452.

Worell, J., & Robinson, D. (1995). Issues in clinical assessment with women. In J. Butcher (Ed.), *Clinical personality assessment. Practical approaches* (pp. 158–171). New York: Oxford University Press.

World Health Organization. (1992). *The ICD-10 classification of mental and behavioural disorders.* Geneva, Switzerland: Author.

Yonkers, K. A., & Gurguis, G. (1995). Gender differences in the prevalence and expression of anxiety disorders. In M. V. Seeman (Ed.), *Gender and psychopathology* (pp. 113–130). Washington, DC: American Psychiatric Press.

Yutzy, S. H., Cloninger, C. R., Guze, S. B., Pribor, E. F., Martin, R. L., Kathol, R. G., Smith, G. R., & Strain, J. J. (1995). DSM–IV field trial: Testing a new proposal for somatization disorder. *American Journal of Psychiatry, 152,* 97–101.

Zahn-Waxler, C. (1993). Warriors and worriers: Gender and psychopathology. *Development and Psychopathology, 5,* 79–89.

Zoccolillo, M. (1993). Gender and the development of conduct disorder. *Development and Psychopathology, 5,* 65–78.

Zoccolillo, M., & Rogers, K. (1991). Characteristics and outcome of hospitalized adolescent girls with conduct disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 30,* 973–981.

Received July 18, 1996
Revision received August 25, 1997
Accepted September 9, 1997 ■

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.