Assessment and Diagnosis of Personality Disorder: Perennial Issues and an Emerging Reconceptualization

Lee Anna Clark

Department of Psychology, University of Iowa, Iowa City, Iowa 52242; email: la-clark@uiowa.edu

Key Words
personality disorder, assessment, comorbidity, diagnosis, stability

Abstract
This chapter reviews recent (2000–2005) personality disorder (PD) research, focusing on three major domains: assessment, comorbidity, and stability. (a) Substantial evidence has accrued favoring dimensional over categorical conceptualization of PD, and the five-factor model of personality is prominent as an integrating framework. Future directions include assessing dysfunction separately from traits and learning to utilize collateral information. (b) To address the pervasiveness and extent of comorbidity, researchers have begun to move beyond studying overlapping pairs or small sets of disorders and are developing broader, more integrated common-factor models that cross the Axis I–Axis II boundary. (c) Studies of PD stability have converged on the finding that PD features include both more acute, dysfunctional behaviors that resolve in relatively short periods, and maladaptive temperamental traits that are relatively more stable—similar to normal-range personality traits—with increasing stability until after 50 years of age. A new model for assessing PD—and perhaps all psychopathology—emerges from integrating these interrelated reconceptualizations.
INTRODUCTION

Personality disorder (PD) is defined in the Diagnostic and Statistical Manual of Mental Disorders—Fourth Edition, Text Revision (DSM-IV-TR; Am. Psychiatr. Assoc. 2000) as sets of traits (stable, longstanding, and pervasive patterns of affectivity, cognition, interpersonal functioning, and impulse control with onset by early adulthood) that are inflexible and maladaptive, deviate markedly from cultural expectations, and cause either significant functional impairment or subjective distress. Based generally on this conceptualization, knowledge about PD has ballooned in the past 20 years, and in a recent editorial in the American Journal of Psychiatry, Gabbard (2005) proclaimed that PD had “come of age” (p. 833).

However, conceptual difficulties and controversies have persisted (Clark et al. 1997, Livesley 2003, Widiger & Samuel 2005), there is widespread dissatisfaction in the field, and articles critiquing the domain are common (Jablensky 2002, Livesley 2003, Millon 2002). Widiger et al. (2002) declared, “Official diagnoses are substantially arbitrary, often unreliable, overlapping, and incomplete and have only a limited utility for treatment planning” (p. 435), and Tyrer et al. (2006) stated bluntly, “The assessment of personality disorder is currently inaccurate, largely unreliable, frequently wrong, and in need of improvement.”

However, with the turn of the century, the field also seems to have turned a corner, as research findings have (a) led to new assessment approaches based on convincing evidence that the structure of PD is dimensional (Trull & Durrett 2005), (b) compelled researchers to think more deeply about the theoretical implications of PD comorbidity both within and between axes (Krueger 2005), and (c) challenged a simplistic view of PD as unchanging (Clark 2005a, Shea & Yen 2003, Tyrer et al. 2006). Together, these results are moving the field toward a more sophisticated understanding of PD and its relation to Axis I disorders. In this context, this review’s goals are (a) to analyze how recent (particularly 2000–2005) research is gradually clarifying these three domains—assessment, comorbidity, and stability; (b) to describe the growing consensus in each; and (c) to articulate the reconceptualization that is emerging and the

PD: personality disorder

DSM: diagnostic and statistical manual

Comorbidity: co-occurrence of two diagnoses in an individual

Contents

INTRODUCTION ........................................... 228
TOPICS NOT REVIEWED ................. 229
HOW SHALL I ASSESS THEE?
LET ME COUNT THE WAYS . . . . . 229
TOWARD DIMENSIONS AND
AWAY FROM CATEGORIES IN
DSM-V? ....................................... 229
Personality Disorder Diagnosis and
the FFM ...................................... 230
Alternative Conceptualizations .... 232
Assessing Dysfunction ................. 234
Alternative Sources of Information 235
Focusing Future PD Assessment
Research ...................................... 236
COMORBIDITY .................................. 236
Complications of Comorbidity ....... 237
Comorbidity with Depression and
Anxiety Disorders ................. 237
Comorbidity with Substance Abuse
and Addictive Behavior .......... 238
Comorbidity with Eating
Disorders ............................... 239
The Meaning of Comorbidity
Patterns .................................... 239
STABILITY .................................... 240
Stability Revisited: Set in Clay, not
Like Plaster ............................. 240
Further Considerations ............. 242
RECONCEPTUALIZING
PERSONALITY DISORDER ....... 244
New Models for Assessing PD .... 244
Toward a Unified Model of Axis I–
Axis II ................................. 246
major directions in which future research should focus.

**TOPICS NOT REVIEWED**

This review is necessarily focused, so I simply mention here three broad topics and five specific ones not covered, to give the reader a sense of the breadth and depth of PD research. Psychopathy research has run parallel with that on DSM antisocial PD, and not only has survived, but has thrived, with active literatures on its assessment, structure, core nature—including relations with antisocial PD and behavior—and universality across age, gender, and culture. Schizotypy research often is biologically oriented due to its relation to schizophrenia, examining that disorder’s full range of neurocognitive deficits; also, several studies have investigated the structure of self-reported dimensions of schizotypy. Focal assessment topics that deserve mention include gender bias, ethnicity/cultural issues, depressive PD (as a possible addition), taxometric analyses [somewhat supportive of schizotypal and antisocial PD categories; mixed, but more suggesting dimensional constructs for borderline personality disorder (BPD), narcissistic PD, and psychopathy; Edens et al. 2006, Fossati et al. 2005, Haslam 2003, Rothschild et al. 2003, Vasey et al. 2005], and the utility of brief screening instruments.

BPD is the most widely researched single-PD domain (Blashfield & Intoccia 2000), encompassing biological, psychosocial, and cognitive factors in etiology and maintenance, functioning, and health care utilization. Related research examines relevant dimensions, including impulsivity, aggression, affective dyscontrol, suicidality, dissociation, traumatic memories and attentional control, as well as attachment and parental rearing style. Because BPD is more likely dimensional than taxonic, and because of poor convergence between BPD measures (Clark et al. 1997), these latter approaches likely will yield more fruitful results.

**HOW SHALL I ASSESS THEE? LET ME COUNT THE WAYS**

Progress in conceptualizing PD continues to be hampered by limitations in its assessment. Simultaneously, improvement in assessing PD is limited by inadequate conceptualization (Clark et al. 1997). In this section, I discuss the current state of—and emerging directions in—both assessment approaches and related conceptual issues, including planning for DSM-V, the role of the five-factor model (FFM) of personality in shaping the field, increasing interest in the critical component of dysfunction, and nonself-report-based assessment.

**TOWARD DIMENSIONS AND AWAY FROM CATEGORIES IN DSM-V?**

A late-2004 research-planning workshop sponsored jointly by the American Psychiatric Association, World Health Organization, National Institute of Mental Health, National Institute on Alcohol Abuse and Alcoholism, and National Institute on Drug Abuse focused on models of PD, both what is known and what needs to be determined for a dimensional model to be adopted in DSM-V. The presented papers and ensuing discussion examined dimensional models from multiple perspectives, from behavior genetics and neurobiology to childhood antecedents and cultural factors, from alternative dimensional structures to clinical utility (Widiger et al. 2005), but the overarching theme and consensus were clear: The current categorical system is scientifically untenable (Widiger & Simonsen 2005), and as Allen Frances declared almost 15 years ago with regard to implementing a dimensional system for PD in the DSM, “Not whether, but when and which” (Frances 1993, p. 110).

Focusing on Frances’s “which,” Widiger & Simonsen (2005) described 18 candidate dimensional systems for PD. Arguing that simply selecting one is scientifically unacceptable...
N/NA: neuroticism/negative affectivity

Five-factor model of personality: posits that five broad, higher-order dimensions of personality exist: neuroticism (N), extraversion (E), agreeableness (A), conscientiousness (C), and openness to experience/culture (O) because each has strengths and weaknesses, they presented a common integrative, hierarchical model topped by two “superfactors” (essentially Digman’s 1997 alpha and beta factors or Block’s 2001 ego resiliency and control), with middle layers of three to seven broad dimensions, which in turn are composed of facets (basic personality traits), all anchored at yet lower levels in specific affects, behaviors, and cognitions. Such a model has its detractors (Block 2001) and limitations in describing personality comprehensively (Hooker & McAdams 2003), and would need further specification to have clinical utility (Verheul 2005), but the common model presented emerges from substantial research and represents a solid base around which to develop a reliable, valid, and scientifically and clinically useful dimensional system.

Often-voiced concerns about dimensional systems regard whether and how they can inform clinical decision-making, and that their complexity complicates clinical communication. Verheul (2005), however, turned the complexity argument around, noting that dimensions provide diagnostic richness and subtlety not afforded by the DSM categories, which are criticized for oversimplicity. Moreover, he argued that the current diagnostic system does not direct either treatment selection or planning; rather, severity is the primary determinant of the decision to treat (see also Tyrer 2005). Moreover, dimensions provide more information for predicting the effectiveness of different treatment options at both the “macro” (e.g., in- versus outpatient, session frequency/duration) and “micro” (e.g., targeting self-harm for initial intervention) levels (Verheul 2005, pp. 293–294). Trull (2005) further discussed the need to determine the most appropriate cut scores on relevant dimensions for various clinical decisions. For example, empirically developed, nonarbitrary cutpoints on trait aggression or self-harm could guide clinical decisions to implement anger-management training or a therapeutic contract regarding suicidal intentions, respectively. Even those who argue for retaining categories (Paris 2005) acknowledge that they will be replaced eventually, most likely by a still developing comprehensive and consensual dimensional system: “Not whether, but when and which” (Frances 1993, p. 110).

Personality Disorder Diagnosis and the FFM

In the last half of the twentieth century, researchers of normal-range personality made tremendous progress in understanding trait structure. By the turn of the century, there was widespread (though not universal; Block 2001) agreement that the Big Five—neuroticism (negative affectivity/emotionality/temperament; N/NA), extraversion (positive affectivity/emotionality/temperament; E/PA), agreeableness (A), conscientiousness (C), and culture/openness to experience (O)—that is, a five-factor model (FFM)—reflected the bulk of personality trait variance. Widiger & Simonsen (2005) tapped the FFM as the framework for organizing the 18 extant dimensional models of PD. Research studies examining relations between the DSM PDs and measures of normal-range personality, including the FFM, have revealed that the domains of normal and abnormal personality are largely overlapping (O’Connor 2002). More specifically, the DSM PDs can be characterized with the FFM conceptually—by both clinical researchers (Widiger et al. 2002) and practicing clinicians (Samuel & Widiger 2004, 2006; Sprock 2002, 2003)—and empirically (O’Connor 2005, Saulsman & Page 2004).

Whereas extant normal-range FFM measures may be limited in differentiating among individuals with more severely maladaptive traits, the FFM per se is not, which has been shown using questionnaire items (Haigler & Widiger 2001), semistructured interviews (Bagby et al. 2005a), and adjective descriptors (Coker et al. 2002). Furthermore, the FFM has been related to PD in translation (De Clercq & De Fruyt 2003); in adolescent
samples (De Clercq & De Fruyt 2003, Lynam et al. 2005); as conceptualized in the ICD-10 (Brieger et al. 2000); and in relation to other psychopathologies including psychopathy (Lynam et al. 2005), depressive PD (Bagby et al. 2004, Huprich 2003), and dependency (Bornstein & Cecero 2000). Lower-order (facet-level) characterizations differentiate better among individuals with PD than do higher-order (domain-level) ones (Bagby et al. 2005a, Morey et al. 2002, Reynolds & Clark 2001). Moreover, practicing clinicians rated facets as more useful clinically than domain scores (Sprock 2002).

Clinicians preferred the FFM to the DSM for describing actual cases and did so reliably (Samuel & Widiger 2006). When rating vignettes of prototypic and nonprototypic cases, interrater reliability was acceptable for all cases using the FFM, but only for prototypic cases (which are rare in actuality) using categories (Sprock 2003). Yet, clinicians’ confidence in their (unreliable) diagnostic and (reliable) FFM ratings of the nonprototypic cases was nearly identical. Thus, the FFM—especially the facet level—appears to have broadband applicability in assessing PD-relevant traits, as well as superior psychometric properties. However, education to familiarize clinicians with using dimensions to diagnose PD will be important when a dimensional system is eventually implemented.

Prototype method of diagnosis. Both to demonstrate that the FFM “possesses the language necessary for the description of the personality disorders” (p. 402) and to utilize clinicians’ familiarity with the DSM diagnoses, Lynam & Widiger (2001) developed a prototype method of diagnosing PD with the FFM. Specifically, expert clinical researchers rated “the prototypic case” (p. 403) of each DSM-IV PD using the 30 facets of the Revised NEO Personality Inventory (NEO PI–R; Costa & McCrae 1992). Their ratings were aggregated, yielding an FFM profile for each PD, which diagnoses PD using similarity scores, intraclass coefficients that assess how closely an individual’s FFM profile matches the prototype profiles.

Miller and colleagues (2004) found good agreement ($r = 0.75$) between these expert prototypes and actual FFM facet-PD score correlations using the Structured Interview for DSM-IV Personality (SIDP-IV; Pfohl et al. 1995) to diagnose PD. Similarity scores were stable (median $rs \geq 0.80$) at 6- and 12-month follow-up, and similarity scores based on the NEO-PI-R and Structured Interview for the Five-Factor Model (SFFM; Trull et al. 2001) converged (median $r = 0.68$) (Miller et al. 2005b).

Miller et al. (2004, 2005b) then correlated FFM similarity scores and actual PD scores in the three samples, obtaining moderate convergence (mean $rs$ in the 0.40s; ranges = 0.02–0.68) using the SIDP-IV or Structured Clinical Interview for DSM-IV Personality Disorders Questionnaire (SCID-II; First et al. 1997) to diagnose PD, and both the NEO-PI-R and SFFM to assess the FFM. Ranges were anchored consistently by obsessive-compulsive PD (OCPD) and avoidant PD (low and high $rs$, respectively). Higher mean convergence ($r = 0.64$) was found with the Schedule for Nonadaptive and Adaptive Personality (SNAP; Clark 1993) diagnostic scale scores, perhaps due in part to shared, self-report method variance. Predictive validity of intake similarity scores for consensus diagnoses based on DSM PD criterion ratings was modest: range = 0.33 (intake) to 0.44 (one year later). Finally, Miller et al. (2005a) demonstrated that a simple sum of component FFM facets performed as well as the prototype method (average convergence = 0.39).

Notably, the convergence of similarity and interview-based scores in these five samples is quite similar to the 0.39 reported in a meta-analysis of self-report and interview convergence (Clark et al. 1997). Thus, these results support the FFM’s claim as a contender in PD assessment, but also demonstrate that the FFM does not surpass the typically moderate convergence between PD measures.
Furthermore, it is critically important to emphasize that diagnosis-by-prototype is not itself an end—the DSM diagnoses are much too flawed to warrant emulation, especially only moderately convergent emulation. The FFM has great value in PD assessment, but it lies in the dimensions themselves and their potential for deepening our understanding of PD traits, not in their ability to approximate demonstrably inadequate categories. The field will be little advanced by additional studies using this approach, its purpose already having been fulfilled.

**Limitations and implications for future directions.** These findings indicate that a lower-level FFM personality trait structure could supplant the current categorical system for diagnosing PD, but that—just as with extant measures—certain limitations must be overcome. First, the NEO PI-R is the only existing faceted FFM measure, so the extent to which its particular facets comprehensively cover and validly represent lower-order levels of the FFM domains is unknown. Moreover, even when facets are used to predict DSM PD, as noted with profile-similarity data, substantial unexplained variance remains. Recent studies using multiple regression, maximizing predictive power, confirm this finding: Bagby et al. (2005a), Reynolds & Clark (2001), and Furnam & Crump (2005) all reported moderate $R^2$s with various combinations of self-report and interview for assessing PD and personality.

Importantly, this limitation is not unique to FFM measures: The SNAP had incremental predictive power over the FFM (mean $\Delta R^2 = 0.22$ and 0.08 for domain and facet scores, respectively), but the reverse also was true (NEO-PI-R scores’ mean $\Delta R^2 = 0.04$ and 0.10 for domains and facets, respectively) (Reynolds & Clark 2001). When three questionnaires were used to predict PD scores, each had incremental predictive power over the others (Bagby et al. 2005b). Conversely, Trobst et al. (2004) found the NEO-PI-R had widely varying predictive power depending on how PD was assessed.

The point is not the limitations of these measures for predicting DSM PD per se—as stated above, those diagnoses are too flawed to be a gold standard. Rather, the diagnoses encompass important clinical problems that comprehensively valid PD measures should assess. Trull (2005) terms this “coverage”—the extent to which a model or system of personality pathology adequately represents those conditions or symptoms that are frequently encountered by clinicians and studied by psychopathologists” (p. 263)—and frames his discussion in terms of content and construct validity.

Thus, an important challenge facing the DSM-V PD work group will be determining how best to capitalize on the strengths of existing measures of the overall personality-PD space—both those developed specifically to assess either PD diagnoses or traits, and those that target normal-range traits with relevance to PD—to provide a maximally comprehensive yet efficient assessment of adaptive and maladaptive personality traits. Put concretely, whereas the FFM generally characterizes the PD domain space as well as do measures specifically designed for that purpose [e.g., the Dimensional Assessment of Personality Pathology–Basic Questionnaire (DAPP-BQ; Livesley & Jackson 2006) and SNAP], other measures also contribute unique, clinically important variance. Because it is impractical to assess PD traits by administering multiple dimensional-system interviews and questionnaires just to ensure that as much valid variance as possible is tapped, identifying and including this additional variance in more comprehensive future PD trait measures is critical to increase both validity and clinical utility of PD-domain assessment.

**Alternative Conceptualizations**

**SWAP-200.** The Shedler-Westen Assessment Procedure (SWAP-200; Westen & Shedler 1999a) is a Q-sort procedure designed
to quantify the richness of clinical description. Westen & Shedler (1999b) reported it had seven factors, which they matched to six DSM PDs (paranoid, schizoid, antisocial, histrionic, narcissistic, and obsessive) plus dysphoric, the largest factor, which yielded five subfactors that both matched (avoidant) and did not match (hostile-external) DSM categories. They also used the SWAP-200 to create DSM prototypes and a high-functioning prototype (Westen & Shedler 1999a). A later factor analysis of the same item pool, however, yielded 12 “clinically relevant personality dimensions” (Shedler & Westen 2004, p. 1743), but neither its relation to the 7-factor analysis nor the motivation for refactoring the instrument was discussed. Moreover, many items marked different factors (e.g., “antisocial-psychopathic PD” in the 7-factor, but “schizoid orientation” in the 12-factor solution) or marked a factor in one solution and no factor in the other. Furthermore, an adolescent version yielded a partially overlapping set of “11 dimensions of adolescent personality” (Westen et al. 2005, p. 227), the Big Five can be found in a subset of SWAP-200 items (Shedler & Westen 2004), and the 11 adolescent dimensions relate systematically to a brief adjective measure of the FFM (Westen et al. 2005).

However, it is not clear how these sets of SWAP-200 factors are related. For example, the 12-factor solution’s “dissociation” dimension does not appear elsewhere; “histrionic sexualization” in the 11- and 12-factor solutions may map onto the DSM and/or factor-analytic histrionic diagnoses, but these empirical relations are not reported. Although a few small-sample studies have reported promising results with the seven factors (Diener & Hilsenroth 2004) or DSM prototypes (Martin-Avellan et al. 2005), additional studies of the 12 factors have not been reported. Accordingly, whether the SWAP-200 structure is robust or relatively sample-dependent, or why different structures have emerged, is unknown.

Currently, therefore, the contribution of the SWAP-200 lies in its demonstration that clinical language and judgment can provide useful, elaborated, and systematic descriptions of the PD space, which underscores the earlier observation of the importance of identifying and including in future PD measures the PD-relevant variance that may not be well represented in extant FFM and other measures.

**Temperament and Character Inventory.**

Cloninger (1987) proposed a theoretical model linking three “temperament” dimensions (harm avoidance, novelty seeking, and reward dependence, from which persistence was split off later) to underlying neural substrates and specific PDs. Tests of the model have yielded mixed results: Predictions regarding the relations between PD and the Temperament and Character Inventory (TCI) have received more support than those concerning neural substrates (Mitropoulou et al. 2003, Mulder & Joyce 2002). Cloninger et al. (1993) added three “character” dimensions (self-directedness, cooperativeness, and self-transcendence), said to develop from experience, to the three (or four) “innate” temperament dimensions.

Low self-directedness consistently marks a wide range of psychopathology, not limited to PD, as does low cooperativeness, although less consistently and less strongly (Daneluzzo et al. 2005, Mulder et al. 1999). Harm avoidance marks subjective distress, and its correlation pattern parallels that of FFM N/NA, being most strongly and consistently high in cluster C (avoidant, dependent, and OCPD) (Farabaugh et al. 2005, Maggini et al. 2000, Mulder et al. 1999). Novelty seeking, like FFM A and C, marks “externalizing” disorders (Krueger et al. 2002, 2005) including substance abuse (Ball 2004, Fassino et al. 2004) and cluster B (antisocial, borderline, histrionic, and narcissistic) PD (Farabaugh et al. 2005, Mulder et al. 1999). Reward dependence is associated with cluster A (paranoid, schizoid, schizotypal) PD in some studies (Farabaugh et al. 2005, Mulder et al. 1999),
as is self-transcendence with schizotypal PD (Daneluzzo et al. 2005).

The TCI is one of the more widely used measures outside the United States, so it has the potential to test the cross-cultural generalizability of PD-relevant constructs. However, O'Connor (2002) investigated the structural robustness of widely used personality and psychopathology measures across clinical and nonclinical samples. Most measures exhibited structural invariance across sample types, whereas the TCI had unstable factor structures both across and within sample types, which may explain its literature’s inconsistencies. Nonetheless, the instrument also may contain clinically relevant variance that a comprehensive PD-domain measure should assess.

Assessing Dysfunction

Like all psychopathology, PD requires either subjective distress or functional impairment, but researchers paid comparatively little attention to the latter until recently. Hill and colleagues (2000) examined the extent to which trait abnormality was separable from dysfunction, and concluded that their two measures assessed similar constructs. Johnson and colleagues (2000b) reported that PD at baseline predicted levels of—and increases in—interpersonal and global dysfunction at one-year follow-up, controlling for HIV status and Axis I disorders. An epidemiological study (Hong et al. 2005) found that all but two PD dimensional scores predicted global functioning 13–18 years later, and half still did so controlling for current Axis I disorder. Another study (Johnson et al. 2005) reported that dysfunction was as strongly predicted by PD—not otherwise specified (PD-NOS) as it was by any other PD. Two reviews found PD was associated with reduced quality of life (Narud & Dahl 2002) and dysfunction “in nearly every realm of concern to healthcare providers” (Smith & Benjamin 2002, p. 135).

The Collaborative Longitudinal Personality Disorders Study (CLPS) examined functioning in four PD groups and in no-PD depressed controls. At intake, patients with schizotypal and BPD had significantly poorer functioning in social relationships and at work and recreation than did those with OCPD or depression only; functioning of those with avoidant PD was intermediate (Skodol et al. 2002). However, at two-year follow-up, despite decreases in PD symptomatology (Shea et al. 2002, Shea & Yen 2003), significant functional improvement had occurred in only three of seven domains—spouse/partner relationships, recreation, and global social adjustment—and that was due largely to changes in the depressed-only group (Skodol et al. 2005d). Notably, patients with BPD or OCPD had no change in dysfunction over the two-year period, except those whose BPD symptoms had improved during the first year. CLPS researchers (Skodol et al. 2005d) concluded that functional impairment might be a more enduring component of PD than the diagnostic criteria per se.

Furthermore, CLPS researchers (Skodol et al. 2005c) examined functioning in patients who had baseline PD and three-year follow-up depression, and compared those with and without persistent PD to depressed patients seen as part of a Medical Outcomes Study (MOS). Functioning was highest in the MOS group, intermediate in those with depression and baseline-only PD, and lowest in those with persistent PD. Dimensional diagnostic scores correlated more strongly with functioning measured at baseline than did scores on higher-order dimensions—either the FFM or the three maladaptive dimensions (negative temperament, positive temperament, and disinhibition versus constraint) of the SNAP (Skodol et al. 2005b), but the SNAP’s lower-order dimensions of maladaptive personality consistently predicted functioning more strongly, up to four-year follow-up (Morey et al. 2006).

However, Mulder (2002) reviewed the literature relating personality and PD to treatment outcome in depression and found mixed results, with the least PD effect in the
best-designed studies. Nonetheless, high FFM-N/NA consistently predicted worse outcome, and PD was never related to better response. Mulder also noted that depressed patients with PD may receive less adequate treatment in some studies. Putting these findings together, it appears that PD has both current and longer-lasting effects on functioning (related CLPS findings: Grilo et al. 2000, 2005); some of these effects may be through relations with Axis I disorders (Hong et al. 2005), and the effects may be attenuated with adequate treatment (Bajaj & Tyrer 2005, Reich 2003).

Noting that assessment of criteria and functioning often are confounded in PD diagnoses, Parker and colleagues (2002) advocated separating their assessment. They reviewed the literature to identify markers of dysfunction, consolidated them into 17 constructs, and developed a 67-item self-report measure of dysfunction with 11 intercorrelated scales representing 2 correlated ($r = 0.64$) higher-order dimensions, noncoping and noncooperativeness (Parker et al. 2004). Factor scales (created from the 10 highest-loading items on each) correlated moderately ($r = 0.56$), and both primary and factor scales correlated moderately (mean $r = 0.50$) with dimensional scores of DSM and ICD-10 diagnoses, suggesting either that measures based on current PD-category concepts tap combinations of traits and dysfunction, or that extremes of personality dimensions are inherently dysfunctional. Research testing these two possibilities is needed.

Self-reported dysfunction correlated with ratings made by a close informant ($r = 0.61$ and 0.41 for noncoping and noncooperativeness, respectively) but not with clinician ratings of functioning across five domains ($rs = 0.16–0.18$). Ratings by two independent interviewers correlated strongly with each other ($rs = 0.70$s), but not with either patient ($rs = 0.33–0.34$) or informant ($rs = 0.25–0.36$) ratings, suggesting that valid ratings of dysfunction may require in vivo life experience with a person (Parker et al. 2004; see also Milton et al. 2005). Testing dysfunction and PD scores together in a regression analysis, noncoping and self-defeating PD scores best differentiated patients from controls, with no other scales adding predictive power (Parker et al. 2004). These results are encouraging steps toward a two-component diagnostic process, but validation samples were small and the findings need replication.

**Alternative Sources of Information**

Consideration of who (or what) can provide the most reliable and valid information for assessing PD began no later than the middle 1980s and remains an important open question. Recent research has transcended simple studies of agreement—which typically is modest and variable (Klonsky et al. 2002, Walters et al. 2004)—to ask more sophisticated questions, such as what factors influence self-informant agreement and who is the best informant for what information.

A review of 17 PD self-informant studies (Klonsky et al. 2002) found no systematic variance by assessment instrument type (questionnaire versus interview), sample (patient versus nonclinical), or informant (except higher agreement with spouses), nor did selves or informants report more overall pathology. Interinformant agreement is generally higher than self-informant agreement, which varies positively by age, degree of acquaintance, and trait/disorder “visibility” (e.g., aggression and BPD are more visible than mistrust and narcissistic PD), and negatively with sample size. Agreement is higher for dimensional ratings and personality problems assessed outside the DSM framework, perhaps because DSM disorders are less internally consistent.

Oltmanns & Turkheimer (2006) summarized multiple studies systematically exploring self-informant agreement. Important findings include: (a) when asked to rate themselves as they thought others would rate them (meta-perceptions), self-informant agreement increased; (b) target- and investigator-selected informants provided similar ratings, although
Clusters A, B, C; paranoid, schizoid, and schizotypal PDs comprise Cluster A; antisocial, borderline, histrionic, and narcissistic PDs comprise Cluster B; avoidant, dependent, and obsessive-compulsive PDs comprise Cluster C.

target-selected informants underrated traits associated with clusters A and B and over-rated cluster C traits; (c) selves and informants provided complementary data (for example, self-rated suspiciousness was rated as "cold" by informants, whereas other-rated suspiciousness was rated as "angry" by selves); and (d) self-ratings of internalizing problems and informant ratings of externalizing problems each added independent predictive power to functional outcomes. Furthermore, when interview- and informant-based ratings were discrepant on narcissism, blind raters of nonverbal behavior in videotaped interviews showed increased agreement with informants. Thus, interviewer ratings may be more accurate if based on all behavior—not just verbal responses—although unreliability is a concern and must be examined.

Utilizing written-record data is a long-standing tradition in psychopathy research, but otherwise has been used rarely in PD assessment. Tyrer et al. (2006) demonstrated that most traits could be rated from records as reliably as from interviews, and convergence with consensus diagnoses was promising in a pilot study. Based on a comprehensive review, Meyer (2002) found that diagnostic agreement correlated with the extent of measures' source-information overlap, and multiple-source measurement had greater reliability and validity than did single-source measurement.

In sum, although not without their own weaknesses, other-source data have distinct advantages over self-reports. Moreover, informant studies have revealed how self-reports may be improved. The field should consider seriously how not only informant data but also meta-perceptions, nonverbal behavior, and written records might be incorporated systematically into PD assessment.

**Focusing Future PD Assessment Research**

Adoption of a dimensional conceptualization of PD inevitably lies ahead, and several key issues must be addressed to arrive at that future. First, the higher-order structure of adaptive and maladaptive personality is well mapped, but lower-order structure remains largely uncharted territory. Because the higher-order dimensions are too broad to capture personality's rich complexity, they have limited utility in clinical settings, and better understanding of more focal traits is critically needed. Furthermore, PD diagnosis is incomplete if only traits are assessed: Personality function also must be evaluated. Extreme personality traits are linked empirically with dysfunction and may be inherently maladaptive, but to investigate this issue, we need dysfunction measures not (or at least less) confounded by personality-trait content. Thus, exploring the nature of dysfunction also should be a field priority.

Finally, it is clear that personality—both adaptive and maladaptive—is too complex to be assessed fully from a single perspective. More comprehensive understanding of PD will require integration of the common and unique information that can be provided by self-report, well-known informants, clinicians who have an objective view contextualized by a broad understanding of the PD landscape, written records of behavior, and eventually laboratory data. Learning how information from these various sources can be integrated most validly and usefully likely will challenge researchers for some years to come.

**COMORBIDITY**

Comorbidity—with its implication of co-occurring but independent disorders—is now widely recognized as a misnomer for the pervasive phenomenon of two or more mental disorders co-occurring; however, the mislabel has stuck and I, too, use it. PD comorbidity has been investigated so much that one would think the topic exhausted, but it is such a fundamental issue that PD-comorbidity research is still increasing. A simple PsycINFO search crossing “personality disorder(s)” and “comorbidity/co-occurrence” yielded more
than 1500 citations from 1985 through 2005: 269 in the first 10 years, 477 in the second five years, and 756 beginning in 2000. Part of the increase stems from an ever-widening sphere of investigation revealing PD comorbidity with, for example, ADHD (Davids & Gastpar 2005) and kleptomania (Grant 2004).

Many studies are purely descriptive, reporting the rates of comorbidity within Axis II and/or with Axis I disorders (typically 50% or more in each case in clinical samples), with discussion of methodological factors that may increase rates artifically, common features, or putative underlying shared etiology [Zimmerman et al. (2005) reported on a particularly large outpatient sample]. Most see PD comorbidity as a problem for the current categorical system, some as its nemesis: “A categorical approach to PDs, resulting in a list of diagnoses, appears useless in psychiatric practice” (Marinangeli et al. 2000, p. 69). The high prevalence of PD-NOS (Verheul & Widiger 2004) may reflect the same fundamental phenomenon as comorbidity—that personality pathology is rarely confined to a single diagnostic entity. Therefore, I focus on new findings and perspectives that add substantially to our understanding. I begin with general consequences of comorbidity, examine results involving particular disorders, and conclude with implications for nosology.

Complications of Comorbidity

Whereas there may be increased prevalence of comorbid PD in samples seeking treatment for Axis I disorders, comorbidity is sufficiently prevalent in population studies that selection bias alone cannot account for it. There is a strong association of PD comorbidity with earlier age of onset (Brieger et al. 2002, Ozkan & Altindag 2005); greater clinical severity (Ozkan & Altindag 2005); poorer treatment outcome (Farabaugh et al. 2005, Ogrodniczuk et al. 2001); longer time to remission (Grilo et al. 2005, Massion et al. 2002); lower long-term social, cognitive, and occupational functioning (Bank & Silk 2001, Denys et al. 2004, Smith & Benjamin 2002, Tyrer et al. 2003); greater medical utilization (Smith & Benjamin 2002); suicide attempts and completion (Garno et al. 2005, Hawton et al. 2003); and greater risk of psychopathology in offspring (Abela et al. 2005). However, worse outcome—including increased time to remission—is not inevitable (Grilo et al. 2000), varies by PD (Grilo et al. 2005), and to some extent may reflect methodological flaws rather than true effects (Mulder 2002).

Comorbidity with Depression and Anxiety Disorders

Two reviews and the results of several large studies paint a coherent picture of relations among depression and anxiety disorders and maladaptive personality. The National Epidemiologic Survey on Alcohol and Related Conditions (NESARC) assessed more than 43,000 individuals and found both pervasive comorbidity among PDs (Grant et al. 2005a) and PDs (especially avoidant and dependent PD) with mood and anxiety disorders (Grant et al. 2005b). Moreover, anxiety/depression and PD each predict the onset of the other (Goodwin et al. 2005, Hettema et al. 2003) and share genetic variance both with each other and with trait N/NA (Bienvenu & Stein 2003). These relations are unaffected by comorbidity with substance use disorders (Verheul et al. 2000).

A 12-year longitudinal study showed that “cothymia” (mixed anxiety-depressive symptoms), higher baseline self-reported anxiety/depression, plus premorbid PD predicted the least favorable outcome, whereas initial Axis I diagnosis per se did not predict outcome, and “instability of [Axis I] diagnosis over time was much more common than consistency” (Tyler et al. 2004, p. 1385). Tyler et al. (2003) suggest that splitting the traditional category of neurosis into specific disorders [e.g., generalized anxiety disorder (GAD), social phobia, dysthymia] was neither helpful nor warranted because of their
extensive comorbidity. They argue provocatively that their results demonstrate “a failed classification system,” that much comorbidity is really “consanguinity,” and that current diagnostic concepts have little impact and “are relatively useless” (p. 136).

Within this picture of general overlap, specific patterns do occur. For example, two studies comparing PD comorbidity in unipolar versus bipolar depression converged on the finding that BPD, OCPD, and avoidant/dependent PD were more common in unipolar patients, and narcissistic PD was more common in bipolar patients (Brieger et al. 2003, Schiavone et al. 2004). Smith et al. (2005), however, found BPD features—especially suicidality and impulse-anger dyscontrol—were more characteristic of bipolar disorder. BPD also has been linked to PTSD (Axelrod et al. 2005). Avoidant/dependent PD predicted reduced remission in GAD (both PDs) and social phobia (avoidant PD only) but not in panic disorder (Massion et al. 2002). Furthermore, some anxiety disorders may be heterogeneous, with different subtypes having differential relations with PD. For example, Lee & Telch (2005) divided OCD obsessions into autonomous obsessions—highly aversive and threatening (e.g., sexual, aggressive, blasphemous, or repulsive) thoughts, images, or impulses—and reactive obsessions—thoughts, doubts, or concerns in which the perceived threat is the possible negative consequence (e.g., of contamination, mistakes, accidents, asymmetry, or disarray). Only the former was associated with schizotypal PD.

### Comorbidity with Substance Abuse and Addictive Behavior

The general picture emerging from PD/substance abuse comorbidity research has been fairly consistent over the years: (a) The overlap is strong, especially with cluster B PDs (Ball 2005). (b) Severity moderates comorbidity; for example, comorbidity with antisocial PD increases about twofold (~30% to 60%) from mild to severe drug abuse/dependence (Flynn et al. 1996). (c) Comorbidity is associated with worse treatment outcome that may be attenuated with enhanced treatment, and substance abusers with PD have earlier substance use, have more legal and family problems, and are more susceptible to relapse in the presence of cravings, negative physical and emotional states, and interpersonal conflict (Ball 2005, Westermeyer & Thuras 2005). (d) Substance abuse and antisocial PD likely share a genetically based etiological factor (e.g., Jang et al. 2000, Kendler et al. 2003, Krueger et al. 2002)—temperamental “externalizing”—which likely also underlies other cluster B PDs, particularly BPD (Bornvalova et al. 2005, Trull et al. 2000) and narcissistic PD (Kelsey et al. 2001), as well as pathological gambling (Pietrzak & Petry 2005) and ADHD (Dowson et al. 2004). This broad factor is characterized in large part by impulsivity, a term that has come under increasing scrutiny as evidence has accrued that it is used to denote several unrelated constructs, some measured by self-report and others via laboratory tasks, each of which has some research support (Bornvalova et al. 2005, Looper & Paris 2000, Whiteside & Lynam 2001). (e) Suicidality and other expressions of negative affectivity may also play a role in PD/substance use comorbidity, although not uniquely so (Bornvalova et al. 2005, Casillas & Clark 2002).

Although much research is conducted on various disorders in the externalizing spectrum from biological perspectives, there is little research, other than behavior genetic studies, examining substance abuse/PD relations from these perspectives, perhaps in part because of limitations in PD measurement, and also because research has tended to be disorder centered (e.g., substance abuse research, BPD research). With growing awareness that diverse externalizing disorders share common biological substrates, we may expect an increase in integrated studies from cognitive-affective neuroscience, and
Comorbidity with Eating Disorders

Eating disorders (EDs)—particularly anorexia nervosa (AN) and bulimia nervosa (BN)—comprise the only diagnostic group besides anxiety/depression and externalizing disorders with a sizeable PD-comorbidity literature. A meta-analysis (Sansone et al. 2005) of four ED groups—AN-restricting (AN-R), AN-bingeing (AN-B), BN, and binge eating disorder (BED)—revealed interesting comorbidity patterns with PDs in all three clusters. Avoidant PD was comorbid with all types of ED, whereas BN showed the strongest comorbidity across the PD domain. Of note, the highest comorbidity rate for both AN-R and BED was with OCPD, and BED also overlapped with cluster A PDs more than other EDs. Severe eating disorders with bingeing were associated with BPD, and BN also overlapped with antisocial and narcissistic PD more than did other EDs. Finally, PD traits were largely absent in obese patients (van Hanswijck de Jonge et al. 2003).

A recent review of personality traits and EDs (Cassin & von Ranson 2005) reported generally complementary findings: Both AN and BN were associated with avoidant PD traits, plus traits that commonly characterize most PDs (e.g., FFM N/NA), whereas AN and BN related to opposite ends of a constraint-disinhibition dimension (see Favaro et al. 2005 for a detailed study of impulsive behaviors in ED). Interestingly, BN as well as AN patients score high on perfectionism and related OC traits (Halmi 2005), which suggests that OC traits are not simply facets of constraint, as is often assumed.

These findings provide a generally coherent picture of ED-PD overlap, but in a special issue of Eating Disorders: The Journal of Treatment & Prevention that focused on ED-PD comorbidity, Vitousek & Stumpf (2005) caution against assessing personality traits and disorders in ED individuals until after the initial treatment phase because of assessment difficulties (e.g., “state” effects of semistarvation and chaotic eating, denial/distortion in self-report, and instability of ED subtypes). Furthermore, CLPS data suggest that ED-PD comorbidity rates reflect disorder base rates rather than meaningful associations (Grilo et al. 2003), and in a general-population twin sample, correlations of PD traits and ED symptoms were modest and nonspecific (Livesley et al. 2005). Thus, observed relations may reflect general associations (e.g., between N/NA and subjective distress) more than specific etiologies, and the large PD/ED comorbidity literature may be “much ado about nothing.”

The Meaning of Comorbidity Patterns

Over the years, writers have offered various theoretical possibilities for explaining comorbidity (Clark 2005b), but rarely have these hypotheses been tested directly. Recently, researchers have begun to analyze comorbidity data in ways that can inform psychopathological theory. For example, regarding the validity of the Axis I–Axis II distinction, CLPS data have revealed both convergences and disjunctions between concurrent and longitudinal co-occurrence (McGlashan et al. 2000, Shea et al. 2004). Specifically, both at baseline and longitudinally up to two years follow-up, avoidant PD was associated with social phobia, and BPD with PTSD (Axelrod et al. 2005 also found bidirectional BPD-PTSD relations in combat veterans). However, BPD/substance abuse and avoidant PD/OCD associations were observed only at baseline, whereas a specific BPD/depression association was found longitudinally, but not at baseline. Others who have found concurrent BPD/depression relations (Bellino et al. 2005) typically have not examined association specificity. OCPD and OCD were not associated concurrently or longitudinally.
Concurrent, but not longitudinal, associations suggest artifactual overlap (e.g., due to shared criteria, such as impulsive substance use, or “state” common factors), whereas longitudinal associations more likely reflect shared underlying pathological structures or processes. For example, Klein & Schwartz (2002) tested several etiologic models and found that a common factor best accounted for longitudinal associations between BPD and early onset dysthymia.

Clark (2005b) analyzed parallels between the literatures examining (a) comorbidity both within-axis and between-axes and (b) relations of personality dimensions to both Axis I and Axis II disorders. Regarding comorbidity, its pervasiveness and extent demonstrated the need to move beyond study of overlapping pairs or small sets of disorders to a broader, more integrated focus. Regarding personality-disorder relations, the data indicated, surprisingly, that personality traits do not have a “privileged” relation with PD, but are equivalently correlated with Axis I and II disorders. Accordingly, Clark (2005b) outlined a general framework to explain both comorbidity and personality-psychopathology relations, suggesting that both domains have common roots in basic temperamental dimensions.

This view is highly congruent with that of Rothbart & Posner (2006), whose comprehensive review of the temperament and developmental psychopathology literature—including their own central contributions linking temperament and attention to neural networks—discusses how temperament and environment act separately and in combination to increase or decrease risk for psychopathology. Relatedly, taxonomies of personality and trait-related symptoms in children and adolescents both strongly suggest they are precursors of the adult FFM dimensions (e.g., De Clercq et al. 2006, Mervielde et al. 2005, Shiner 2005). This work should be required reading for all PD researchers.

**STABILITY**

A definitional assumption has been that PD is enduring. Some have argued that stability is the sole province of PD, that the inclusion on Axis I of “early onset, chronic impairments that characterize everyday functioning” and the absence of a clear distinction between the two types of disorder reflects inadequate conceptualization ( Widiger 2003, p. 90). Moreover, questions about PD diagnostic stability began as early as 1985 ( Barasch et al. 1985), when a longitudinal BPD study reported a stability of 77%, termed “relatively stable” (p. 1486). However, they did not compute the statistic kappa ( )—new at that time—which was a low-moderate 0.46.

Dimensional PD criterion counts proved more stable than diagnoses ( Loranger et al. 1991), and Zimmerman’s (1994) review found average $\kappa$ = 0.56 for both short- and long-term retest stabilities for “any PD.” These findings implicated measurement error in diagnostic instability, with underlying personality pathology more stable. That is, if no more diagnostic change occurs over longer compared with shorter intervals, then the observed change likely is artifactual, due to measurement error (e.g., resulting from minor change across diagnostic boundaries). Otherwise, longer interval coefficients should be lower, reflecting either greater measurement error over longer time spans or both error and true change. Interestingly, for individual diagnoses, shorter interval studies did have higher $\kappa$s ( Zimmerman 1994).

**Stability Revisited: Set in Clay, not Like Plaster**

Recently, findings from four major studies and several more focused ones have stimulated reconsideration of PD stability. The CLPS reported significant diagnostic and criterion-level change over two years (Grilo et al. 2004, Shea et al. 2002): Only 44% of patients met criteria every month during year one, and...
baseline-to-two-years \( \kappa \) (corrected for rater unreliability) were in the low 0.50s. Remission rates (fewer than two criteria throughout the past year) averaged around 20% at year one and ranged from 20% to 40% at year two. Mean criterion levels for their four target diagnoses dropped, from baseline, an average of 22% (six months), 33% (one year), and 41% (two years).

The Longitudinal Study of Personality Disorders (LSPD), a four-year college student study, found statistically significant mean decreases of 1.4 PD criteria per year, ranging from near 0 in women without baseline PD to just over 2.5 in males with baseline PD, with considerable individual variation (Lenzenweger et al. 2004). The large Children in the Community (CIC) study, examining PD traits three times from an average age of 14 to 22, found steady declines, with overall PD trait levels decreasing 28% (Johnson et al. 2000a).

Finally, Durbin & Klein (2006) reported on a follow-along sample of depressed outpatients who were assessed every 30 months for 10 years. The median \( \kappa \) for any PD meeting full criteria for the four 2.5-year periods, the three 5-year periods, the two 7.5-year periods, and the full 10 years were 0.34, 0.47, 0.42, and 0.23, whereas including cases with one criterion fewer than threshold yielded somewhat higher \( \kappa \)s of 0.51, 0.37, 0.46, and 0.29, which suggests the arbitrariness of the DSM’s diagnostic thresholds. Specific PD diagnoses were notably more unstable. It is unclear whether the lower 10-year values represent lasting or transient change, given the notably higher 7.5-year stabilities.

The McLean Study of Adult Development (MSAD), an intensive, longitudinal study of BPD, reported similar change rates and patterns (Zanarini et al. 2005). Two smaller studies examined two-year PD stability in adolescents. One obtained similar results (significantly lower scores at reassessment; Grilo et al. 2001), whereas the other found diagnostic stability only in antisocial PD, and moderate stability in dimensional PD scores, with no overall decline (Chanen et al. 2004).

Despite diagnostic and criterial instability, all these studies found notably stronger rank-order stability of criterion counts. For example, even given the very restricted range of the DSM PD’s 7- to 9-point scales, and the fact that all patients were above diagnostic threshold at baseline (limiting range still further), the CLPS stability coefficients in the PD groups from baseline were 0.74 (six months), 0.67 (one year), and 0.59 (two years) (Grilo et al. 2004). For the total sample, the six-month and one-year stabilities were impressively high: 0.90 and 0.86, respectively (Shea et al. 2002). Similarly, two- to three-year and nine-year rank-order stabilities of any PD in the CIC were 0.69 and 0.52, respectively (Johnson et al. 2000a), whereas average two-year stability in another adolescent sample was 0.50 (Chanen et al. 2004). Durbin & Klein (2006) reported a median intraclass coefficient of 0.59 over 10 years for the three clusters, and 0.49 for specific PDs [range = 0.23 (antisocial) to 0.61 (avoidant)].

How are we to interpret these data? For example, the CLPS met the DSM general diagnostic criteria for PD, including having a long-standing, enduring trait pattern. The CLPS protocol specifically examined “the prior two years, but traits must be characteristic of the person for most of his or her adult life in order to be counted toward a diagnosis” (Shea et al. 2002, p. 2037). Because of this two-year window, if patients were judged not to have enduring patterns of behavior at one-year follow-up, then traits judged at baseline to be enduring either were validly judged enduring at baseline but were not evident at all during the ensuing year, or were not validly judged at either baseline or one-year follow-up—indicating measurement error. That is, either patients changed their story at reassessment and/or interviewers changed their judgment about traits being characteristic of the patients’ adult lives and also manifest in the past two years.
In considering these possibilities, given that the DSM definition of personality disorder builds on personality traits, PD stability and change must be evaluated in the context of a broader literature. In 1980, when DSM-III introduced a separate axis for PD, there were few good data on personality stability, and the prevailing notion was that of William James (1950/1890) who claimed that “by the age of 30, the character has set like plaster, and will never soften again” (p. 121). In the intervening 25 years, however, research has challenged this notion, and a more sophisticated understanding of normal-range personality stability and change has emerged.

The main findings can be summarized briefly: Mean-level trait change is moderate through adolescence and well into early adulthood, with increasing levels of positive traits (e.g., A and C) and decreasing levels of negative traits (e.g., N/NA) (Robins et al. 2001, Vaidya et al. 2002). Furthermore, characteristic affect levels are less stable than broad personality dimensions, at least in part because they have stronger relations with positive and negative life events (Vaidya et al. 2002). Rank-order stability is moderate in childhood ($M = 0.31$) and continues to increase through adolescence and adulthood, not peaking ($M = 0.74$) until considerably later, between ages 50 and 70 (Roberts & DelVecchio 2000). Personality structure is quite stable as early as adolescence, whereas individual profile stability is moderate, with more change in level and scatter (spread) than in profile shape (Robins et al. 2004, Vaidya et al. 2002). Thus, change in basic personality configuration is more quantitative than qualitative, and quantitative change is not insignificant.

Given these normative data, there may be yet another possibility in evaluating the observed stability/change in PD besides simply true change and measurement error, and that is that interviewers applied normative standards to their judgments of whether a trait is enduring and characteristic, with the result that the observed level of PD stability closely parallels that of normal traits. For example, CLPS two-year criterion stability in specific diagnoses for patients (initial $M_{age} = 33$) is 0.59, whereas normal-range trait stability for ages 30–39 is 0.62 (Roberts & DelVecchio 2000). Durbin & Klein (2006) provided a more direct comparison by assessing personality with the Eysenck Personality Questionnaire-Revised (EPQ-R; Eysenck et al. 1985). In comparison with the median 0.59 for PD scored dimensionally, median stability for personality was highly similar: 0.65. Thus, these recent studies do not actually challenge the definition of PD traits as stable and longstanding. Rather, they inform us that our intuitive sense of personality stability corresponds to lower numbers than we might have predicted a priori, which concerns the level of our ability to quantify subjective experience rather than the phenomenon per se. Nonetheless, modifying James, it isn’t until past the age of 50 that character may set like plaster; before then, it’s more like being set in clay—change can occur, but gradually and with effort.

**Further Considerations**

There are several other issues to consider before drawing final conclusions about PD stability. First is the important theoretical question of whether maladaptive personality may be expected to have more, less, or the same stability as normal-range personality. PD is defined as inflexible, possibly suggesting more stability, but this term likely is intended instead to indicate lack of situational adaptivity. On the other hand, given the strong affective component in most PD (Trobst et al. 2004, Widiger et al. 2002) and the lower stability of trait affect, PD may be less stable than normal-range personality. This issue deserves further consideration, but absent compelling reason to hypothesize otherwise, it is reasonable to assume that the empirical results are face valid—PD has comparable stability to normal-range personality.

Second, recalling that PD diagnoses are less stable than their component traits, it also
is important to consider PD stability in the context of Axis I disorder stability. A comparison of remission rates from four longitudinal studies found PD more stable than mood disorders, but generally less so than anxiety disorders (Shea & Yen 2003). Median 1- and 2-year remission rates, respectively, were 0.78 and 0.87 for mood disorders, 0.31 and 0.45 for PD, and 0.14 and 0.20 for anxiety disorders (except 0.40/0.60 in nonagoraphobic panic). Similarly, the Nottingham Study of Neurotic Disorder (Tyrer et al. 2004) found considerable specific-diagnosis instability, but only 36% overall remission after 12 years in patients with GAD, panic disorder, or dysthymia (either with—cothymia—or without comorbid anxiety disorder). Remission was even lower in patients with cothymia and/or PD. Given the close correspondence of anxiety disorder and N/NA, basic temperament dimensions likely undergird the observed diagnostic stability. Durbin & Klein (2006), on the other hand, found comparable stabilities between PD and anxiety disorders overall (10-year \( \kappa \) for any anxiety disorder = 0.24), with panic disorder and obsessive-compulsive disorders more stable (10-year \( \kappa \) = 0.32 and 0.39, respectively) than social or simple phobia (0.07 and −0.02, respectively).

Third, recall that CLPS researchers found high two-year stability of maladaptive functioning (Skodol et al. 2005c; see also Tyrer et al. 2004). Taken together, these three points indicate that whereas PD features (i.e., traits) have the same stability as normal-range personality and PD diagnoses are only moderately stable, PD-based dysfunctional behavior is relatively more stable, which may contribute to “the commonly held belief that PDs persist” (Skodol et al. 2005d, p. 444). That is, persistent dysfunction may be interpreted as a persistent diagnosis.

Fourth, as noted almost 15 years ago (Shea 1992), the PD diagnostic criteria are qualitatively variable. Some (e.g., sense of entitlement) directly assess their target construct(s) (i.e., a particular enduring trait pattern defining that PD), whereas others (e.g., uncomfortable or helpless when alone) assess not the target construct itself (attachment, in this example), but rather “manifestations of an underlying construct” (Shea 1992, p. 37). Among the latter, a further distinction is the degree to which criteria tap acute, dysfunctional behaviors that more closely resemble Axis I symptoms (e.g., recurrent suicidal behavior) that resolve in shorter periods than do maladaptive traits per se (Clark 2005a, Shea et al. 2002).

In the CLPS, examples of more changeable criteria were odd behavior and constricted affect (schizotypal PD), self-injury and behaviors to avoid abandonment (BPD), avoiding interpersonal jobs and potentially embarrassing situations (avoidant PD), and miserly and strict moral behaviors (OCPD).

The MSAD (Zanarini et al. 2005) provided additional evidence of this distinction, linking BPD’s moderately high remission rates to acute symptoms (e.g., suicidal/self-mutilating behaviors) that diminished with treatment, maturation, or stress resolution, whereas basic temperament dimensions (e.g., chronic anger, stress reactivity) changed more slowly—perhaps at the rate of normal-range affective trait change, which is less stable still than nonaffective traits.

Fifth, given that dimensional PD scores are more stable than diagnoses and that PDs are composed of traits (manifested in specific criteria), how do personality traits relate to PD diagnoses when assessed independently of PD diagnoses? In the CLPS, latent longitudinal models demonstrated significant cross-lagged relations between year-one FFM traits and year-two PD diagnoses (but not baseline—year one) for three PDs (not OCPD), beyond the considerable trait and PD diagnostic (rank-order) stability, whereas the reverse was not observed (Warner et al.
Thus, change in personality traits predicted PD change, but not vice versa.

The lack of the effect from baseline to year one likely is because early PD change primarily reflects amelioration of the Axis I–like symptoms that led the patient to seek treatment, rather than trait-level change. The longitudinal associations of certain Axis I disorders with PD discussed above provide some support for this hypothesis. In particular, all significant change occurred between baseline and month 6, with no significant change between 6 and 12 months for any of the PD groups (Shea et al. 2004; see also Grilo et al. 2005).

Sixth and finally, different amounts of information are available at baseline, 6, and 12 months, which means that they are necessarily somewhat different assessment methods with increasing reliability and, concomitantly, validity, based on the principle of aggregation. Specifically, at baseline, only retrospective information is available, whereas at later assessments, 6 or 12 months of prospective data also are available. Assuming even modest consistency, aggregation of more observations increases measurement reliability, with a decreasing rate of change as more and more observations accrue (thus, greater change in the first versus second 6 months of observations). This suggests that initial single-point-in-time PD assessment should be considered “provisional,” and that a “definite” diagnosis of PD requires multimodal (self- and informant-report and documentary) evidence of temporal duration of individuals’ particular personality trait sets.

RECONCEPTUALIZING PERSONALITY DISORDER
The picture that emerges from joint consideration of these issues is one of change within relative stability. More specifically, personality and PD reflect similarly structured trait combinations (O’Connor 2002, 2005) and have moderate long-term stability, but differ in extremity, maladaptivity, and consequential behavior. Importantly, PD is defined further by acute symptoms that are linked directly to maladaptive traits (e.g., avoiding interpersonal occupational activity with social inhibition), and/or develop as defensive or compensatory behaviors (e.g., self-mutilation) to cope with stress—both exogenous and self-created by one’s own maladaptivity (Skodol et al. 2005a). These more changeable symptoms, together with the inherent lesser reliability of single-observation assessment, largely account for observed diagnostic instability, while patients’ personality traits change at the “normal” rate, and being more extreme and maladaptive, account for persistent dysfunction.

This emerging view of PD is consistent with (a) the current view of personality as reflecting longstanding—but not immutable—characteristics that are based in genetic inheritance as well as both early and ongoing life experiences (Caspi et al. 2005) and (b) empirical data showing that PD diagnoses are, at most, moderately stable. Importantly, this model suggests that PD diagnosis should be approached differently from the current symptom/criterion method, distinguishing assessment of more acute symptoms from that of patients’ basic temperament.

Additionally, in this conceptualization, traditional, single-point-in-time and single-source-of-information assessment cannot and should not be expected to yield entirely valid PD diagnoses. We know this intuitively with regard to normal personality, understanding that first impressions only partially capture a person’s true nature, that to know someone well, we must interact with them in a variety of situations, discussing a range of topics. PD is no different. The very nature of personality and PD demands longitudinal and multimodal assessment for validity.

New Models for Assessing PD
Reconceptualizing PD must be accompanied by new assessment models; four strikingly similar ones have been offered. Livesley et al.
(1994), ahead of the field, first proposed that all disorders be on Axis I, with personality traits on Axis II. PD diagnosis would have two components (Livesley & Jang 2000). (a) Diagnosing disorder—determining dysfunction and assessing acute dysfunctional behaviors (e.g., ideas of reference, aggression, suicidality, hyperperfectionism). Diagnosis would entail determining personality’s functional failure to solve major life tasks: developing an integrated sense of self, adaptive interpersonal relationships, and adaptive social functioning. (b) Describing the individual’s personality traits, organized in four (FFM minus O) broad dimensions (Livesley 2005), but with particular focus on lower-order dimensions. Parker et al. (2000) replicated Livesley’s dimensions with an independent set of descriptors and, as noted above, reported promising results for a two-component (dysfunction, traits) diagnostic process (Parker et al. 2002, 2004). Both Livesley and Parker drew parallels between their models of functioning and Freud’s definition of psychological health, “lieben und arbeiten” (to love and to work).

Widiger et al. (2002) proposed a four-stage model: Assess (a) individuals’ higher- and lower-order FFM personality profiles, (b) specific impairments secondary to extreme traits, (c) whether dysfunction is clinically significant (e.g., GAF score), and (d) if possible, profile match to a particular PD prototype. They emphasized that some situations might not require all four steps (e.g., a counseling setting where formal diagnosis was unnecessary). Tyrer et al. (2006) advocated assessing (a) clinically important personality dysfunction severity (versus PD per se) (Tyrer 2005) and (b) four broad dimensions (also FFM minus O), with greater reliability than the DSM’s narrower categories and more historical continuity than three or five dimensions, (c) using multiple information sources, including written records, and (d) longitudinal assessment.

These models differ in order, emphasis, and detail, and each has particular strengths: Livesley’s conceptualization has the most theoretical emphasis, linking a PD diagnosis to the functional failure of personality to solve the major life tasks for which it is designed (Wakefield 2006). Parker et al. (2000) developed a promising measure of functioning. Widiger et al.’s model is the most well specified, providing both an established assessment instrument and detailed descriptions of potential problems associated with each facet’s extremes. Tyrer et al’s (2006) model affords the broadest assessment, incorporating multiple information sources and longitudinal assessment.

The models’ similarities, however—particularly distinguishing dysfunction and trait assessment, and essential agreement on four of five broad dimensions—are far more important than their differences, and an integrated model could incorporate each models’ strengths: A diagnosis of PD—focused on personality-based dysfunctionality—would be recorded on Axis I, whereas individuals’ relevant personality traits, that is, the relatively stable characteristics that underlie the Axis I PD diagnosis, would be recorded on Axis II. Initially, diagnosis would be provisional, pending confirmation of the assessed traits’ stability and their validity via multiple-source assessment. Rather than the DSM/ICD system, which requires meeting a small number of criteria (currently a mixture of trait manifestations and trait-dysfunction blends), existing psychometrically sound trait assessment measures would be used, and new ones developed, taking advantage of advances in personality theory and assessment, including understandings of personality structure and process (Mroczek & Cooper 2006), cross-cultural generalizability of personality constructs (Ashton & Lee 2005), informant measures (Oltmanns & Turkheimer 2006), written records (Tyrer et al. 2006), and computer adaptive testing (Simms & Clark 2005). Thus, PD trait assessment could evolve as the field matured and could incorporate clinical utility as a development criterion. Assessment of dysfunction needs much more development, but promising starts have been made.

Widiger et al. (2002) proposed a four-stage model: Assess (a) individuals’ higher- and lower-order FFM personality profiles, (b) specific impairments secondary to extreme traits, (c) whether dysfunction is clinically significant (e.g., GAF score), and (d) if possible, profile match to a particular PD prototype. They emphasized that some situations might not require all four steps (e.g., a counseling setting where formal diagnosis was unnecessary). Tyrer et al. (2006) advocated assessing (a) clinically important personality dysfunction severity (versus PD per se) (Tyrer 2005) and (b) four broad dimensions (also FFM minus O), with greater reliability than the DSM’s narrower categories and more historical continuity than three or five dimensions, (c) using multiple information sources, including written records, and (d) longitudinal assessment.

These models differ in order, emphasis, and detail, and each has particular strengths: Livesley’s conceptualization has the most theoretical emphasis, linking a PD diagnosis to the functional failure of personality to solve the major life tasks for which it is designed (Wakefield 2006). Parker et al. (2000) developed a promising measure of functioning. Widiger et al.’s model is the most well specified, providing both an established assessment instrument and detailed descriptions of potential problems associated with each facet’s extremes. Tyrer et al’s (2006) model affords the broadest assessment, incorporating multiple information sources and longitudinal assessment.

The models’ similarities, however—particularly distinguishing dysfunction and trait assessment, and essential agreement on four of five broad dimensions—are far more important than their differences, and an integrated model could incorporate each models’ strengths: A diagnosis of PD—focused on personality-based dysfunctionality—would be recorded on Axis I, whereas individuals’ relevant personality traits, that is, the relatively stable characteristics that underlie the Axis I PD diagnosis, would be recorded on Axis II. Initially, diagnosis would be provisional, pending confirmation of the assessed traits’ stability and their validity via multiple-source assessment. Rather than the DSM/ICD system, which requires meeting a small number of criteria (currently a mixture of trait manifestations and trait-dysfunction blends), existing psychometrically sound trait assessment measures would be used, and new ones developed, taking advantage of advances in personality theory and assessment, including understandings of personality structure and process (Mroczek & Cooper 2006), cross-cultural generalizability of personality constructs (Ashton & Lee 2005), informant measures (Oltmanns & Turkheimer 2006), written records (Tyrer et al. 2006), and computer adaptive testing (Simms & Clark 2005). Thus, PD trait assessment could evolve as the field matured and could incorporate clinical utility as a development criterion. Assessment of dysfunction needs much more development, but promising starts have been made.
This model has theoretical implications for treatment as well. For example, it suggests that dialectical behavior therapy (Robins et al. 2004) may succeed, at least in part, by sequencing its treatment objectives, first targeting Axis I behaviors that are life threatening, interfere with treatment, and lower quality of life, and then shifting focus to Axis II problems—developing adaptive life skills (e.g., anger management, conflict resolution) and resolving longstanding problematic interpersonal dynamics.

**Toward a Unified Model of Axis I–Axis II**

No doubt the 1980 separation of PD onto its own Axis II was an important stimulus for launching the current PD research enterprise, and substantial progress has been made in the past 25 years through research largely based on the categorical system in the DSMs. However, as findings cumulate, the validity of the current Axis I–Axis II separation is questioned increasingly. Krueger (2005) called for “a unified model of personality, personality disorders, and clinical disorders” (p. 233); Clark (2005b) argued that basic temperament dimensions provide the basis for personality as well as the development of both Axis I and II disorders, proposing, therefore, that the field work toward a single, hierarchical, integrated framework that would bridge the three domains. Both interdisciplinary work (Rothbart & Posner 2006) and multiple research lines focusing on dimensions of psychopathology (e.g., impulsivity/aggression, sleep disturbance, negative affect, and serotonin dysfunction) without regard to specific diagnoses signal a changing paradigm.

Thus, the integrated PD assessment model described above may be a first step toward an even broader integration: Axis II dimensions may underlie the dysfunctionality of both PD and traditional diagnostic symptom groups recorded on Axis I. A paradigm shift toward a dysfunction-dimensional conceptualization appears to lie in the near future, first for PD, but then for psychopathology in general, with the potential to revolutionize our entire field (Widiger et al. 2005). I urge the field to follow Robert Frost and take this road less traveled by, for that will make all the difference.

**SUMMARY POINTS**

1. On the basis of the weight of the empirical evidence supporting dimensional approaches to personality disorder diagnosis, serious consideration is being given to switching to a dimensional (from the current categorical) system for Axis II in DSM-V.

2. The five-factor model (FFM) of personality is widely accepted as representing the higher-order structure of both normal and abnormal personality traits. Thus, if a dimensional system is implemented for Axis II in DSM-V, most likely it will be based on some variant of the FFM.

3. Because it is recognized that personality disorder cannot be diagnosed based on extreme personality traits alone, there is increasing interest in assessing personality dysfunction and understanding the link(s) between personality traits and dysfunction.

4. Convergence is modest, at best, between self-report and nonself-report-based assessment of personality and personality disorders. Recent research has documented increased reliability and validity of assessment when multiple sources are used to diagnose personality.

5. Comorbidity both within Axis II and between Axes I and II has been well documented; similarly, personality trait dimensions have been shown to relate moderately...
to strongly with both Axis I and Axis II disorders. Converging lines of evidence suggest that personality and psychopathology may have common roots in basic temperament.

6. A definitional assumption has been that personality and PD are enduring, but recent findings have stimulated reconsideration of stability. Specifically, research has established that normal personality becomes more stable over time until at least age 50, and the observed degree of change of personality pathology, when measured dimensionally, appears to be highly similar to that of normal personality. Furthermore, PD criteria have been shown to be a mix of more stable trait dimensions (e.g., low self-worth) and less stable symptomatic behaviors (e.g., self-injury). The lower stability of PD diagnoses (compared with dimensions) may not be simply an artifact of the DSM categorical system, but also may reflect the inclusion of this latter type of criteria.

7. Four strikingly similar models for diagnosing PD have emerged and can be integrated as follows: a diagnosis of PD, focused on personality dysfunction, would be recorded on Axis I, whereas the relevant personality traits, that is, the relatively stable characteristics underlying the Axis I PD diagnosis, would be recorded on Axis II. Initial diagnosis would be provisional, pending confirmation of the assessed traits’ stability via longitudinal assessment and trait levels through multiple information sources.

8. The described integrated PD assessment model may be a first step toward an even broader integration based on Axis I–Axis II comorbidity patterns. As mentioned in Summary Point 5, abnormal temperament may underlie the dysfunctionality of not only PD but also clinical syndromes traditionally recorded on Axis I. Thus, a paradigm shift toward a two-part conceptualization of psychopathology—dysfunction and dimensional assessment—may lie in the future.

FUTURE ISSUES

1. At some point, perhaps as early as DSM-V, the Axis II PDs will be diagnosed using a dimensional conceptualization, but several key issues remain to be addressed. First, for a dimensional system to have utility in clinical settings, a consensually validated lower-order structure of adaptive and maladaptive personality is needed to capture personality’s rich complexity.

2. PD diagnosis is incomplete if only traits are assessed: Personality functioning also must be evaluated. Extreme personality traits are linked empirically with dysfunction and may be inherently maladaptive, but measures of dysfunction that are not—or at least are less—confounded by personality trait content are needed. Thus, exploring the nature of dysfunction should be a field priority.

3. Finally, adaptive and maladaptive personalities are too complex to be assessed fully from a single perspective. More comprehensive understanding of PD will require integration of the common and unique information that is obtained from self-report, well-known informants, clinicians who have an objective view contextualized by a broad understanding of the PD landscape, written records of behavior, and, eventually, laboratory data. Integrating information from these various sources for a fuller understanding of PD will challenge researchers for some years to come.
ACKNOWLEDGMENTS

I wish to thank (a) Leigh Wensman for her tireless and ever-cheerful help in searching and obtaining hundreds of articles, compiling and organizing their abstracts, and cross-checking all references, (b) the too-many-to-name-individually colleagues who provided chapters, impress papers, and excellent feedback on an earlier draft, and (c) Tom Widiger, in particular, for catalyzing me to rethink the entire final section.

LITERATURE CITED


Reviews various factors that distinguish Axis I clinical syndromes from Axis II personality disorders; argues that these disorders are more similar than different.

Examines whether personality disorders are distinct from each other and from normal personality; analyzes arguments for placing personality disorders and clinical syndromes on different axes.


Pietrzak RH, Petry NM. 2005. Antisocial personality disorder is associated with increased severity of gambling, medical, drug and psychiatric problems among treatment-seeking pathological gamblers. *Addiction* 100:1183–93


Examines temporal stability of PD in naturalistic longitudinal studies and compares it to that of depression and anxiety. Contributes to reconceptualization of PD.

Examines stability of functional impairment in the four CLPS PD diagnoses over the first two years of the study.


RELATED RESOURCES


Krueger RF, Tackett JL, eds. 2006. Personality and Psychopathology. New York: Guilford


Documents assessment issues that remained unresolved with regard to PD diagnosis, including test-retest reliability and poor convergent validity of PD measures.
Contents

Prefatory

Research on Attention Networks as a Model for the Integration of Psychological Science
Michael I. Posner and Mary K. Rothbart ................................................................. 1

Cognitive Neuroscience

The Representation of Object Concepts in the Brain
Alex Martin .............................................................................................................. 25

Depth, Space, and Motion

Perception of Human Motion
Randolph Blake and Maggie Shiffrar ................................................................. 47

Form Perception (Scene Perception) or Object Recognition

Visual Object Recognition: Do We Know More Now Than We Did 20 Years Ago?
Jessie J. Peissig and Michael J. Tarr ................................................................. 75

Animal Cognition

Causal Cognition in Human and Nonhuman Animals: A Comparative, Critical Review
Derek C. Penn and Daniel J. Povinelli ................................................................. 97

Emotional, Social, and Personality Development

The Development of Coping
Ellen A. Skinner and Melanie J. Zimmer-Gembeck ........................................... 119
Biological and Genetic Processes in Development

The Neurobiology of Stress and Development
Megan Gunnar and Karina Quevedo .................................................. 145

Development in Societal Context

An Interactionist Perspective on the Socioeconomic Context of Human Development
Rand D. Conger and M. Brent Donnellan ........................................... 175

Culture and Mental Health

Race, Race-Based Discrimination, and Health Outcomes Among African Americans
Vickie M. Mays, Susan D. Cochran, and Namdi W. Barnes ...................... 201

Personality Disorders

Assessment and Diagnosis of Personality Disorder: Perennial Issues and an Emerging Reconceptualization
Lee Anna Clark .................................................................................. 227

Social Psychology of Attention, Control, and Automaticity

Social Cognitive Neuroscience: A Review of Core Processes
Matthew D. Lieberman ........................................................................ 259

Inference, Person Perception, Attribution

Partitioning the Domain of Social Inference: Dual Mode and Systems Models and Their Alternatives
Arie W. Kruglanski and Edward Orehek ............................................. 291

Self and Identity

Motivational and Emotional Aspects of the Self
Mark R. Leary .................................................................................... 317

Social Development, Social Personality, Social Motivation, Social Emotion

Moral Emotions and Moral Behavior
June Price Tangney, Jeff Stuewig, and Debra J. Mashek ......................... 345
The Experience of Emotion
Lisa Feldman Barrett, Batja Mesquita, Kevin N. Ochsner,
and James J. Gross ................................................................. 373

Attraction and Close Relationships
The Close Relationships of Lesbian and Gay Men
Letitia Anne Peplau and Adam W. Fingerhut ................................. 405

Small Groups
Ostracism
Kipling D. Williams ................................................................. 425

Personality Processes
The Elaboration of Personal Construct Psychology
Beverly M. Walker and David A. Winter ........................................ 453

Cross-Country or Regional Comparisons
Cross-Cultural Organizational Behavior
Michele J. Gelfand, Miriam Erez, and Zeynep Aycan ................................ 479

Organizational Groups and Teams
Work Group Diversity
Daan van Knippenberg and Michaél C. Schippers .......................... 515

Career Development and Counseling
Work and Vocational Psychology: Theory, Research,
and Applications
Nadya A. Fouad ................................................................. 543

Adjustment to Chronic Diseases and Terminal Illness
Health Psychology: Psychological Adjustment
to Chronic Disease
Annette L. Stanton, Tracey A. Revenson, and Howard Tennen .................. 565
Research Methodology

Mediation Analysis
   David P. MacKinnon, Amanda J. Fairebild, and Matthew S. Fritz ....................... 593

Analysis of Nonlinear Patterns of Change with Random Coefficient Models
   Robert Cudeck and Jeffrey R. Harring .................................................. 615

Indexes

Cumulative Index of Contributing Authors, Volumes 48–58 ......................... 639
Cumulative Index of Chapter Titles, Volumes 48–58 ................................. 644

Errata

An online log of corrections to Annual Review of Psychology chapters (if any, 1997 to the present) may be found at http://psych.annualreviews.org/errata.shtml