

The Shedler–Westen Assessment Procedure (SWAP): Making Personality Diagnosis Clinically Meaningful

Jonathan Shedler

*Department of Psychiatry
University of Colorado Health Sciences Center*

Drew Westen

*Department of Psychology and Department of Psychiatry and Behavioral Sciences
Emory University*

There is a schism between science and practice in understanding and assessing personality. Approaches derived from the research laboratory often strike clinical practitioners as clinically naïve and of dubious clinical relevance. Approaches derived from clinical observation and theory often strike empirical researchers as fanciful speculation. In this article, we describe an approach to personality designed to bridge the science–practice divide. The Shedler–Westen Assessment Procedure (SWAP; Shedler & Westen, 2004a, 2004b; Westen & Shedler, 1999a, 1999b) is an empirically rigorous diagnostic method that preserves the richness and complexity of clinical case description. In this article, we describe its use in diagnosis, case conceptualization, and treatment planning. We review evidence for reliability, validity, and clinical utility. Finally, in the article, we present a system for personality diagnosis, as an alternative to *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., text rev.; American Psychiatric Association, 2000) Axis II, that is empirically grounded, clinically relevant, and practical for routine use in both clinical and research contexts.

One of the greatest challenges facing psychiatry and psychology is the growing schism between science and practice. The schism is especially pronounced in conceptualizing and assessing personality. For most clinical practitioners, personality diagnosis is a task requiring judgment and expertise. Expert clinicians consider a wide range of psychological data, attending not only to what patients say but also to how they say it, and drawing complexly determined inferences from patients' accounts of their lives and relationships, from their manner of interacting with the clinician, and from their own emotional reactions to the patient (Westen & Arkowitz-Westen, 1998).

For example, clinicians tend not to assess lack of empathy, a diagnostic criterion for narcissistic personality disorder (PD), by administering self-report questionnaires or asking patients direct questions (Westen, 1997). (Not only are narcissistic patients unlikely to endorse such items, they may well describe themselves as caring people and wonderful friends.) An initial sign of lack of empathy on the part of the patient is often a subtle sense on the part of the clinician of being interchangeable or replaceable, of being treated as a sounding board rather than as a fellow human being (for

empirical evidence, see Betan, Heim, Conklin, & Westen, 2005; for a clinical discussion, see McWilliams, 1994). The clinician might go on to consider whether she consistently feels this way with this particular patient and whether such feelings are characteristic for her in her role as therapist. The clinician might then become aware that the patient tends to describe others more in terms of the functions they serve or the needs they meet than in terms of who they are as people. She might further consider whether and how these issues dovetail with the facts the patient has provided about his life, with the problems that led him to seek treatment, with information gleaned from family members or other collateral contacts, and so on.

It is just such clinical judgment and inference that many personality researchers eschew. As successive editions of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM) have minimized the role of clinical inference, investigators have increasingly treated personality diagnosis as a technical task of tabulating signs and symptoms with relatively little consideration for how they fit together, the psychological functions they serve, their meanings, the developmental trajectory that gave rise to them, or the present-day

factors that maintain them. Indeed, the diagnostic “gold standard” in PD research is the structured interview. Such assessment methods are designed to achieve interrater reliability by minimizing the role of clinical judgment and substituting standardized questions and decision rules. Indeed, the interviews are typically administered by research assistants or trainees, not by experienced clinicians.

DSM and structured assessment procedures evolved as they have for good reason. Prior to *DSM-III*, psychiatric diagnosis was unsystematic, overly subjective, and of questionable scientific merit. It sometimes revealed more about the clinician’s background and theoretical predilections than it did about the patient’s personality dispositions. Structured assessment methods evolved in the service of science and in reaction against the unsystematic diagnostic methods of the past. In the evolution of personality diagnosis from a largely subjective, clinical enterprise to a largely technical, research-driven enterprise, much has been gained and much has been lost. The solution to the science–practice schism cannot be to turn back the clock and abandon the scientific advances of the past decades. Nor can it be to disregard the cumulative insights of generations of clinical observers. The solution, rather, may be a marriage of the best aspects of clinical observation and empirical rigor.

In this article, we describe the Shedler–Westen Assessment Procedure (SWAP; Shedler & Westen, 1998, 2004a, 2004b; Westen & Shedler, 1999a, 1999b), an approach to personality assessment designed to *harness* clinical judgment and inference rather than eliminate it and combine the best features of the clinical and empirical traditions in personality assessment. It provides a means of assessing personality that is both clinically relevant and empirically rigorous.

In this article, we (a) review problems with the *DSM* diagnostic system for PDs, (b) discuss the challenges of using clinical observation and inference in research, (c) describe the development of the SWAP as a method for systematizing clinical observation, (d) illustrate its use for diagnosis and clinical case conceptualization, (e) review evidence for reliability and validity, and (f) discuss recommendations for revising Axis II for *DSM-V*.

WHY REVISE AXIS II?

The approach to PD diagnosis codified by *DSM* now finds little favor with either clinicians or researchers. There is consensus that *DSM* Axis II requires reconfiguration. Some of the problems with Axis II include the following (see also Clark, 1992; Grove & Tellegen, 1991; Jackson & Livesley, 1995; Livesley, 1995; Livesley & Jackson, 1992; Westen & Shedler, 1999a, 2000; Widiger & Frances, 1985):

1. The diagnostic categories do not rest on a sound empirical foundation and often disagree with findings from cluster and factor analyses (Blais & Norman, 1977;

Clark, 1992; Harkness, 1992; Livesley & Jackson, 1992; Morey, 1988).

2. *DSM* Axis II commits arbitrarily to a categorical diagnostic system. It may be more useful to conceptualize borderline pathology, for example, on a continuum from none through moderate to severe rather than classifying borderline PD as present/absent (Widiger, 1993). This same consideration applies to individual diagnostic criteria. For example, just how little empathy constitutes “lack of empathy?”
3. *DSM* Axis II lacks the capacity to weight criteria that differ in diagnostic importance (Davis, Blashfield, & McElroy, 1993).
4. Comorbidity between PD diagnoses is unacceptably high. Patients who meet criteria for any PD often meet criteria for four to six PDs (Blais & Norman, 1997; Grilo, Sanislow, & McGlashan, 2002; Oldham et al., 1992; Pilkonis et al., 1995; Watson & Sinha, 1998). This suggests lack of discriminant validity of the diagnostic constructs, assessment methods, or both.
5. In attempting to reduce comorbidity, *DSM* work groups have gerrymandered diagnostic categories and criteria, sometimes in ways faithful neither to clinical observation nor empirical data. For example, they excluded lack of empathy and grandiosity from the diagnostic criteria for antisocial PD to minimize comorbidity with narcissistic PD, even though the traits apply to both PDs (Westen & Shedler, 1999a, 1999b; Widiger & Corbitt, 1995).
6. Efforts to define PDs more precisely have led to narrower criterion sets over time, progressively eroding the distinction between PDs (multifaceted syndromes encompassing cognition, affectivity, motivation, interpersonal functioning, and so on) and simple personality traits. The diagnostic criteria for paranoid PD, for example, are essentially redundant indicators of one trait: chronic suspiciousness. The diagnostic criteria no longer describe the multifaceted personality syndrome recognized by most experienced clinicians (Millon, 1990; Millon & Davis, 1997).
7. *DSM* Axis II does not consider personality strengths that might rule out PD diagnoses for some patients. For example, differentiating between a patient with narcissistic PD and a much healthier person with narcissistic personality dynamics may not be a matter of counting symptoms, but of noting whether the patient has such positive qualities as the capacity to love and sustain meaningful relationships characterized by mutual caring and understanding.
8. *DSM* Axis II does not encompass the spectrum of personality pathology clinicians see in practice. Among patients receiving treatment for personality pathology, fewer than 40% can be diagnosed on Axis II (Westen & Arkowitz-Westen, 1998).
9. *DSM* Axis II diagnoses are not as clinically useful as they might be. For example, knowing whether a patient

meets criteria for avoidant PD or dependent PD tells one little about the function of the person's symptoms, which personality processes to target for treatment, or how to treat them.

10. The algorithm used for diagnostic decisions (counting symptoms) diverges from the methods clinicians use—or could plausibly be expected to use—in real-world practice. Cognitive research suggests that clinicians do not make diagnoses by tabulating symptoms. Rather, they gauge the overall “match” between a patient and a cognitive template or prototype of the disorder (i.e., they consider the features of a disorder as a configuration or gestalt), or they apply causal theories that make sense of the interrelations between symptoms (Blashfield, 1985; Cantor & Genero, 1986; Kim & Ahn, 2002; Westen, Heim, Morrison, Patterson, & Campbell, 2002).
11. PD assessment instruments do not meet standards for reliability and validity normally expected in psychological research. Questionnaires and structured interviews show relatively weak convergence with one another and with the longitudinal evaluation using all available data (LEAD) standard (Perry, 1992; Pilkonis et al., 1995; Skodol, Oldham, Rosnick, Kellman, & Hyler, 1991; Spitzer, 1983; Westen, 1997). They also show poor test–retest reliability at intervals greater than 6 weeks (First et al., 1995; Zimmerman, 1994). Poor test–retest reliability is especially problematic given that PDs are by definition enduring and stable over time.¹

Most of the proposed solutions to these problems share the assumption that progress lies in further minimizing the role of the clinician, either by developing increasingly behavioral and less inferential diagnostic criteria or by bypassing the clinician altogether through the use of self-report instruments. These attempted solutions may, however, be part of the problem. By eliminating clinical observation and inference, investigators may inadvertently be eliminating crucial psychological phenomena from consideration (Cousineau & Shedler, 2006; Shedler, Mayman, & Manis, 1993). An alternative to eliminating clinical inference is to harness it for scientific use.

¹Poor test–retest reliability has led some researchers to suggest that PDs are less stable than previously believed. Such an interpretation of the data seems inconsistent with the observations of virtually all clinical theorists. A more viable hypothesis may be that the assessment instruments do not capture core features of personality that are salient to clinicians who treat patients with PDs and know them well. Specifically, the instruments may overemphasize transient behavioral symptoms (such as self-cutting and suicidality in borderline patients, which may emerge only when an attachment relationship is threatened) and underemphasize underlying personality processes that endure over time (such as affect dysregulation and feelings of emptiness and self-loathing in borderline patients).

THE CHALLENGE OF CLINICAL DATA

The problem with clinical observation and inference is not that it is inherently unreliable, as some investigators have assumed (for a discussion and literature review, see Westen & Weinberger, 2004). The problem is that it tends to come in a form that is difficult to study systematically. Rulers measure in inches and scales measure in pounds, but what metric do psychotherapists share? Imagine three clinicians reviewing the same case material. One might describe the patient in terms of schemas and belief systems, another may speak of conditioning history, and the third of conflicts and defenses.

It is not readily apparent whether the hypothetical clinicians can or cannot make similar observations. There are three possibilities: (1) They may be observing the same thing but using different language and metaphor systems to describe it, (2) they may be attending to different aspects of the clinical material, as in the parable of the elephant and the blind men, and (3) they may not be able to make the same observations at all. To determine whether the clinicians can make the same observations and inferences, one must ensure that they speak the same language and attend to the same spectrum of clinical phenomena.

A STANDARD VOCABULARY FOR CASE DESCRIPTION

The SWAP is an assessment instrument designed to provide clinicians of all theoretical orientations with a standard “vocabulary” for case description. The vocabulary consists of 200 statements, each of which may describe a given patient very well, somewhat, or not at all. The clinician describes a patient by ranking or ordering the statements into eight categories, from those that are most descriptive (assigned a value of 7) to those that are not descriptive (assigned a value of 0). Thus, the SWAP yields a score from 0 to 7 for each of 200 personality-descriptive variables. (A Web-based version of the SWAP can be previewed at www.SWAPassessment.org).

The “standard vocabulary” of the SWAP allows clinicians to provide in-depth psychological descriptions of patients in a systematic and quantifiable form and ensures that all clinicians attend to the same spectrum of clinical phenomena (cf. Block, 1961/1978). SWAP statements are written in a manner close to the data (e.g., “Tends to get into power struggles,” or “Is capable of sustaining meaningful relationships characterized by genuine intimacy and caring”), and statements that require inference about internal processes are written in clear, unambiguous language (e.g., “Tends to see own unacceptable feelings or impulses in other people instead of in him/herself”). Writing items in this jargon-free manner minimizes unreliable interpretive leaps and makes the item set useful to clinicians of all theoretical perspectives.

The SWAP is based on the Q-sort method, which requires clinicians to assign each score a specified number of times (i.e., there is a “fixed distribution” of scores). The SWAP distribution is asymmetric, with many items receiving scores of “0” (not descriptive) and progressively fewer items receiving higher scores. The use of a fixed distribution has psychometric advantages and eliminates much of the measurement error or “noise” inherent in standard rating scales.² The method maximizes the opportunity to observe statistical relations where they exist but does not, as some incorrectly believe, artifactually inflate reliability or validity coefficients. Block (1961/1978) has described the psychometric rationale for the Q-sort method in detail; his psychometric conclusions remain unchallenged and we refer the interested reader to his classic text.

The SWAP item set was drawn from a wide range of sources including the clinical literature on PDs written over the past 50 years (e.g., Kernberg, 1975, 1984; Kohut, 1971; Linehan, 1993); Axis II diagnostic criteria included in *DSM-III* through *DSM-IV*; American Psychiatric Association, 1994); selected *DSM* Axis I items that could reflect aspects of personality (e.g., depression and anxiety); research on coping, defense, and affect-regulatory mechanisms (e.g., Perry & Cooper, 1987; Shedler et al., 1993; Vaillant, 1992; Westen, Muderrisoglu, Fowler, Shedler & Koren, 1997); research on interpersonal pathology in patients with PDs (Westen, 1991, Westen, Lohr, Silk, Gold, & Kerber, 1990); research on personality traits in nonclinical populations (e.g., Block, 1971; John, 1990; McCrae & Costa, 1990); research on PDs conducted since the development of Axis II (see Livesley, 1995); pilot interviews in which observers watched videotaped interviews of patients with PDs and described them using earlier versions of the item set; and the clinical experience of the authors.

Most important, the SWAP-200 (the first major edition of the SWAP item set) is the product of a 7-year iterative item revision process that incorporated the feedback of hundreds of clinician-consultants who used earlier versions of the instrument to describe their patients (Shedler & Westen, 1998). We asked each clinician one crucial question: “Were you able to describe the things you consider psychologically important about your patient?” We added, rewrote, and revised items based on this feedback and then asked new clinicians to describe new patients. We repeated this pro-

cess over many iterations until most clinicians could answer “yes” most of the time. A new, revised version of the SWAP item set, the SWAP-II, incorporates the additional feedback of over 2,000 clinicians of all theoretical orientations. The iterative item revision process was designed to ensure the comprehensiveness and clinical relevance of the SWAP item set.

Because the SWAP is jargon free and clinically comprehensive, it has the potential to serve as a language for describing personality pathology that can be used by any skilled clinical observer. Our studies demonstrate that experienced clinicians of all theoretical orientations understand the items and score them reliably. In one study, a nationwide sample of 797 experienced psychologists and psychiatrists of diverse theoretical orientations, who had an average of 18 years practice experience posttraining, used the SWAP-200 to describe patients with personality pathology (Westen & Shedler, 1999a). These experienced clinicians provided similar SWAP-200 descriptions of patients with specific PDs regardless of their theoretical commitments, and fully 72.7% agreed with the statement “I was able to express most of the things I consider important about this patient” (the highest rating category). In a subsequent sample of 1,201 psychologists and psychiatrists who used the SWAP-II, 84% “agreed” or “strongly agreed” with the statement “The SWAP-II allowed me to express the things I consider important about my patient’s personality” (fewer than 5% disagreed). Again, the ratings were unrelated to clinicians’ theoretical orientation. Virtually identical findings were obtained in a national sample of 950 clinicians who used the adolescent version of the instrument, the SWAP-II-A (Westen, Dutra, & Shelder, 2005; Westen, Shedler, Durrett, Glass, & Martens, 2003). We are unaware of other personality item sets that have been evaluated in this manner for *clinical* comprehensiveness and relevance.

CAPTURING CLINICAL NUANCE

Just as researchers tend to be skeptical regarding the scientific usefulness of clinical observation, many clinicians express skepticism that a structured assessment instrument can do justice to the richness and complexity of clinical case description. However, SWAP statements can be combined in patterns to capture a wide range of subtle clinical phenomena and convey meanings that transcend the content of the individual items.

Consider, for example, the SWAP-II item “Tends to be sexually seductive or provocative.” If a patient receives a high score on this item along with high scores on the items “Has an exaggerated sense of self-importance (e.g., feels special, superior, grand, or envied)” and “Seems to treat others primarily as an audience to witness own importance, brilliance, beauty, etc.,” the portrait that begins to emerge is one of a narcissistically organized individual who seeks sexual attention to bolster a self-concept of being special and uniquely

²One way it does so is by ensuring that raters are “calibrated” with one another. Consider the situation with rating scales, in which raters can use any value as often as they wish. Inevitably, certain raters will tend toward extreme values (e.g., values of 0 and 7 on a 0–7 scale) and others will tend toward middle values (e.g., values of 4 and 5). Thus, the scores reflect not only the characteristics of the patients but also the calibration of the raters. The Q-sort method, with its fixed distribution, eliminates this kind of measurement error because all clinicians must assign each score the same number of times. If use of a standard item set gives clinicians a common vocabulary, use of a fixed distribution can be said to give them a “common grammar” (Block, 1961/1978).

desirable. If the same patient also receives high scores on the items “Tends to feel s/he is not his/her true self with others; may feel false or fraudulent” and “Tends to feel s/he is inadequate, inferior, or a failure,” then a more complex psychological portrait begins to emerge. The SWAP items in combination indicate that feelings of grandiosity and inadequacy coexist in the same person, and suggest the hypothesis that grandiosity serves to mask painful feelings of inadequacy. Indeed, this duality may lie at the heart of narcissistic PD for many patients. The ability to describe and quantify psychological conflict and contradiction is a key feature of the SWAP, one that distinguishes it from other dimensional models (which assume that a person can be high or low on a trait but not both).

If the SWAP–II item describing sexual seductiveness is instead combined with the items “Tends to fear s/he will be rejected or abandoned,” “Appears to fear being alone; may go to great lengths to avoid being alone,” and “Tends to be ingratiating or submissive (e.g., consents to things s/he does not want to do, in the hope of getting support or approval),” the portrait that begins to emerge is one of a dependent individual who relies on sexuality as a desperate means of maintaining attachments in the face of feared abandonment.

If the SWAP–II item describing sexual seductiveness is combined with the items “Tends to act impulsively (e.g., acts without forethought or concern for consequences),” “Takes advantage of others; has little investment in moral values (e.g., puts own needs first, uses or exploits people with little regard for their feelings or welfare, etc.),” and “Experiences little or no remorse for harm or injury caused to others,” the portrait that begins to emerge is one of an antisocial individual who exploits others sexually and whose primary concern is gratifying immediate needs.

If the item describing sexual seductiveness is combined with the items “Has a deep sense of inner badness; sees self as damaged, evil, or rotten to the core,” and “Appears to want to ‘punish’ self; creates situations that lead to unhappiness, or actively avoids opportunities for pleasure and gratification,” one might plausibly infer that sexuality plays a role in a larger pattern of self-devaluation and self-abasement (such a person might well become the victim of the antisocial individual described above). These brief examples illustrate how SWAP items can be combined to communicate subtle clinical concepts, and how the same item can convey different meanings depending on the items that surround and contextualize it. We further illustrate this in a later section with a clinical case example.

TREATMENT IMPLICATIONS

The *DSM* diagnostic criteria are largely descriptive, providing little guidance for clinicians trying to understand the meaning and function of the symptoms or how to intervene. For example, *DSM-IV* indicates that borderline patients

are characterized by “a pattern of unstable and intense interpersonal relationships” (*DSM-IV-TR*, 2000, p. 710). The statement may be descriptively accurate, but *why* does the patient have unstable relationships and how can the clinician help? Because the SWAP addresses underlying personality processes that give rise to these characteristics, it suggests some answers.

Consider the following personality process (Item 9 in the SWAP–II): “When upset, has trouble perceiving both positive and negative qualities in the same person at the same time (e.g., may see others in black or white terms, shift suddenly from seeing someone as caring to seeing him/her as malevolent and intentionally hurtful, etc.)” The item describes the phenomenon known to psychodynamic clinicians as *splitting* and to cognitive-behavioral clinicians as *dichotomous thinking*. If the patient’s perceptions of others gyrate between contradictory extremes, it follows that relationships will be unstable. This implies a specific treatment strategy: The therapist will intervene effectively if he or she can help the patient recognize the extremes of thinking and perceive others in a more balanced light. For example, the therapist may observe, “When you’re angry with your partner, it’s hard to keep in mind that there’s anything you like about him. And when you’re feeling close, it seems hard for you to remember that he has any flaws.” Such interventions are designed to develop the patient’s capacity to integrate contradictory perceptions and see others in more complex, modulated, and balanced ways. A recent clinical trial has demonstrated the efficacy of a treatment for borderline PD based on just this type of intervention (Clarkin, Levy, Lenzenwenger, & Kernberg, 2007; Levy, Meehan, Kelly, et al., 2006).

DSM-IV also indicates that borderline patients may have “transient, stress-related paranoid ideation” (p. 710) but leaves one in the dark about why this occurs or how to intervene. Suppose a patient has high scores on the following SWAP–II items: “Is prone to intense anger, out of proportion to the situation at hand” (Item 185) and “Tends to see own unacceptable feelings or impulses in other people instead of in him/herself” (Item 116). The items, considered in combination, suggest a hypothesis about the meaning and function of paranoid ideation: The patient may become paranoid (i.e., see the world as dangerous and hostile) because, in times of intense agitation, the patient sees his or her own hostility wherever he or she looks. (Empirically, these items do emerge in combination for paranoid patients; see the section “Toward *DSM-V*: An Improved Classification of Personality Disorders” below). The treatment implications are clear: The therapist must help the patient develop more adaptive ways to regulate anger.

SWAP DIMENSIONAL DIAGNOSIS

The SWAP scoring algorithms generate a dimensional score for each PD included in the *DSM-IV* (as well as for factor-analytically derived traits and for an alternative set of

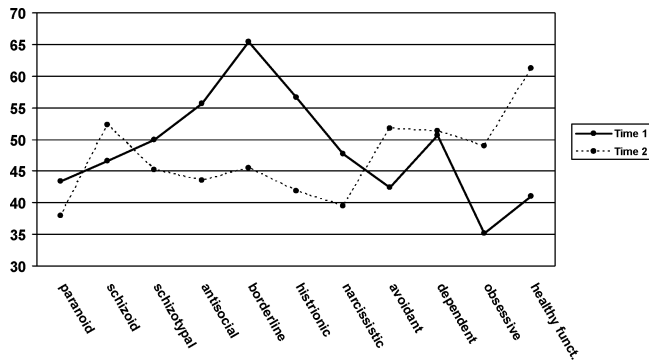


FIGURE 1. The Shedler–Westen Assessment Procedure personality disorder score profile plotted as T scores (normed in a psychiatric sample of patients with Axis II diagnoses). *Note.* From “Assessing Personality Change in Psychotherapy With the SWAP–200: A case study,” by V. Lingardi, J. Shedler, and F. Gazillo, 2006, *Journal of Personality Assessment*, 86, p. 27. Copyright 2006 by Lawrence Erlbaum Associates, Inc. Reprinted with permission.

diagnostic syndromes that we identified empirically; see the section “Toward *DSM–V*: An Improved Classification of Personality Disorders” later). Additionally, the SWAP generates richly detailed narrative case descriptions relevant to clinical case conceptualization and treatment planning. Dimensional PD scores measure the similarity or match between a patient and prototype SWAP descriptions representing each personality syndrome in its typical or “ideal” form (e.g., a prototypical patient with paranoid PD). Dimensional PD scores can be expressed as T scores and graphed to create a PD score profile resembling a Minnesota Multiphasic Personality Inventory (MMPI; Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989) profile as shown in Figure 1. Thus, each personality syndrome is assessed on a continuum rather than diagnosed categorically as present/absent. Low PD scores indicate that the patient does not resemble or match the PD prototype and high scores indicate that the patient matches it well, with intermediate scores indicate varying degrees of resemblance (for descriptions of the PD prototypes and scale construction methods, see Shedler & Westen, 2004b; Westen & Shedler, 1999a, 1999b).

Note that this dimensional approach preserves a syndromal understanding of personality styles and disorders. That is, it treats personality as a configuration of functionally interrelated psychological processes encompassing affectivity, cognition, motivation, interpersonal functioning, coping strategies and defenses, and so forth. By *functionally related*, we mean that the personality processes are interdependent and have causal relations to one another. (e.g., in the example of paranoid ideation given above, intense anger and the propensity to project unacceptable feelings onto others are functionally related). The approach does not deconstruct personality configurations into separate trait dimensions such as those derived from factor analysis of questionnaire data (the approach taken, e.g., by the Five-factor model).

A syndromal approach is consistent with research showing that clinicians view psychopathology in terms of functionally interrelated psychological processes (just as human judgment about category membership more generally relies on implicit causal theories linking component parts into coherent gestalts; Kim & Ahn, 2001). It is also consistent with empirical and conceptual recognitions that personality syndromes fall on continua from relatively healthy through severely disturbed. For example, a relatively healthy person with an obsessional personality style might be precise, orderly, logical, more comfortable with ideas than feelings, a bit more concerned than most with issues of authority and control, and somewhat rigid in certain areas of thought and behavior. Such a person may excel in fields where such attributes are adaptive such as finance, engineering, or perhaps the development of diagnostic systems. Toward the more disturbed end of the obsessional spectrum, we find individuals who are rigidly dogmatic, oblivious to affect, preoccupied with control, and who consistently misapply logic in ways that lead them to miss the forest for the trees.

Although we are emphasizing in this article the utility of a syndromal approach, we do not discount the utility of trait approaches derived from conventional factor analysis. Indeed, factor analysis of the SWAP has identified clinically and empirically coherent trait dimensions (Shedler & Westen, 2004a; Westen, Dutra, & Shedler, 2005; Westen, Waller, et al., 2006), some of which map readily onto trait dimensions included in other dimensional trait models (Widiger & Simonsen, 2005) and some of which do not (e.g., thought disorder, sexual conflict). Syndromal and trait approaches each have advantages for different assessment purposes. A combined approach may well prove most informative, for example, by describing patients syndromally and then adding trait dimensions that are not redundant with the syndromes (e.g., hostility, thought disorder) to create fine-grained psychological portraits.

CASE ILLUSTRATION³

Background

“Melanie” is a 30-year-old White woman with presenting complaints of substance abuse and inability to extricate herself from an emotionally and physically abusive relationship. Assessment with the Structured Clinical Interview for *DSM* (SCID) and SCID–II (First et al., 1995) structured interviews yielded an Axis I diagnosis of substance abuse and an Axis II diagnosis of borderline PD with histrionic traits, with a Global Assessment of Functioning (GAF) score of 45.

³The material in this section is adapted from Lingardi, Shedler, and Gazillo (2006). Please see the original publication for a more complete description of the case, treatment methods, and findings.

Melanie's early family environment was marked by neglect and parental strife. A recurring family scenario is illustrative: Melanie's mother would scream at her husband, telling him he was a failure and that she was going to leave him; she would then slam the door and lock herself in her room, leaving Melanie frightened and in tears. Both parents would then ignore Melanie, often forgetting to feed her. By adolescence, Melanie was often skipping school, spending her days sleeping or wandering the streets. At age 18, she left home and began a "life on the streets," entering a series of impulsive and chaotic sexual relationships, abusing street drugs, and engaging in petty theft. In her mid-20s, Melanie moved in with her boyfriend, a small-time drug dealer. Melanie periodically prostituted herself to obtain money or drugs for her boyfriend, who sometimes beat her when she did not bring home enough.

Melanie began psychotherapy at a frequency of three sessions per week. The first 10 psychotherapy sessions were tape recorded and transcribed. Two clinicians (blind to all other data) reviewed the transcripts and scored the SWAP–200 based on the information contained in the transcripts. The SWAP–200 scores were then averaged across the two clinical judges to enhance reliability and obtain a single SWAP–200 description. After 2 years of psychotherapy, 10 consecutive psychotherapy sessions were again recorded and transcribed, and the SWAP assessment procedure was repeated.

Personality Disorder Diagnosis

The solid line in Figure 1 shows Melanie's PD scores at the beginning of treatment for the 10 PDs included in *DSM–IV*. A "healthy functioning" index is graphed as well, which reflects clinicians' consensual understanding of healthy personality functioning (Westen & Shedler, 1999a). For ease of interpretation, the PD scores have been converted to T scores based on norms established in a psychiatric sample of patients with Axis II diagnoses (Westen & Shedler, 1999a).

Although the SWAP assesses PDs dimensionally and treats personality syndromes as continua, we have also established cutoff scores for "backward compatibility" with the categorical approach of *DSM–IV*. We have suggested $T = 60$ as a threshold for making a categorical Axis II diagnosis, and $T = 55$ as a threshold for diagnosing "features."⁴

Melanie's PD profile shows a marked elevation for borderline PD ($T = 65.4$, approximately $1\frac{1}{2}$ SDs above the sample mean), with secondary elevations for histrionic PD ($T = 56.6$) and antisocial PD ($T = 55.7$). Applying the recommended cutoff scores, Melanie's *DSM–IV* Axis II diagnosis is borderline PD with histrionic and antisocial features. Also noteworthy is the T score of 41 for the "healthy functioning"

index, nearly 1 SD below the mean in a reference sample of patients with Axis II diagnoses. The low score indicates significant impairment in functioning and parallels the low GAF score assigned at intake.

Narrative Case Description

We can generate a narrative case description by listing the SWAP items with the highest scores in the patient's SWAP description (e.g., items with scores of 5, 6, and 7). The narrative description, following, is based on the top 30 most descriptive SWAP–200 items. We have grouped together conceptually related items. To aid the flow of the text, we made some minor grammatical changes and added some summary statements and connecting text (italicized). However, the SWAP–200 items are reproduced essentially verbatim.

Melanie experiences severe depression and dysphoria. She tends to feel unhappy, depressed, or despondent, appears to find little or no pleasure or satisfaction in life's activities, feels life is without meaning, and tends to feel like an outcast or outsider. She tends to feel guilty, and to feel inadequate, inferior, or a failure. Her behavior is often self-defeating and self-destructive. She appears inhibited about pursuing goals or successes, is insufficiently concerned with meeting her own needs, and seems not to feel entitled to get or ask for things she deserves. She appears to want to "punish" herself by creating situations that lead to unhappiness, or actively avoiding opportunities for pleasure and gratification. Specific self-destructive tendencies include getting drawn into and remaining in relationships in which she is emotionally or physically abused, abusing illicit drugs, and acting impulsively and without regard for consequences. She shows little concern for consequences in general.

Melanie shows many personality traits associated specifically with borderline PD. Her relationships are unstable, chaotic, and rapidly changing. She has little empathy and seems unable to understand or respond to others' needs and feelings unless they coincide with her own. Moreover, she tends to confuse her own thoughts, feelings, and personality traits with those of others, and she often acts in such a way as to elicit her own feelings in other people (for example, provoking anger when she herself is angry, or inducing anxiety in others when she herself is anxious).

Melanie expresses contradictory feelings without being disturbed by the inconsistency, and she seems to have little need to reconcile or resolve contradictory ideas. She is prone to see certain others as "all bad," losing the capacity to perceive any positive qualities they may have. She lacks a stable image of who she is or would like to become (e.g., her attitudes, values, goals, and feelings about self are unstable and changing) and she tends to feel empty. *Affect regulation is poor:* She tends to become irrational when strong emotions are stirred up and shows a noticeable decline from her customary level of functioning. She also seems unable to soothe or comfort herself when distressed and requires the involvement of another person to help her regulate affect.

⁴The relatively low thresholds reflect the fact that the reference sample consisted of patients with PD diagnoses. Thus, a T score of 50 indicates average functioning among patients with PD diagnoses, and a T score of 60 represents an elevation of 1 SD relative to other patients with PD diagnoses.

Both her living arrangements and her work life tend to be chaotic and unstable.

Finally, *Melanie's attitudes toward men and sexuality are problematic and conflictual*. She tends to be hostile toward members of the opposite sex (whether consciously or unconsciously) and she associates sexual activity with danger (e.g., injury or punishment). She appears afraid of commitment to a long-term love relationship, instead choosing partners who seem inappropriate in terms of age, status (e.g., social, economic, intellectual), or other factors.

The narrative provides a detailed psychological portrait of a severely troubled patient with borderline personality pathology. The description helps illustrate the difference between descriptive psychiatry (aimed at establishing a diagnosis) and clinical case formulation (aimed at understanding an individual person). In this instance, however, all findings are derived from the same quantitative assessment data.

Melanie's case has a surprisingly happy ending. After 2 years of psychotherapy, the SWAP revealed significant personality change that paralleled objective and considerable changes in Melanie's life circumstances. The dotted line in Figure 1 shows Melanie's PD scores after 2 years of therapy. For a detailed account of the case and the treatment, see Lingardi et al. (2006).

RELIABILITY AND VALIDITY

Psychological and psychiatric researchers often assume that clinical observation and judgment are unreliable; a well-established literature has documented the limitations of clinical judgment. Unfortunately, studies of clinical judgment have too often asked clinicians to make predictions about things that fall outside their legitimate area of expertise (and just as unfortunately, some clinicians have been all too willing to offer such prognostications). More problematic, the studies have typically conflated clinicians' ability to make accurate observations and inferences (which they do well) with their ability to combine and weight variables to derive optimal predictions (a task necessarily performed better by statistical methods such as regression equations). In fact, a substantial literature has documented the reliability and validity of clinical observation and inference when it is quantified and utilized appropriately (Westen & Weinberger, 2004). The SWAP differs from other assessment approaches in that it harnesses clinical judgment using psychometric methods developed specifically for this purpose and then applies statistical and actuarial methods to the resulting data. In short, it relies on clinicians to do what they do best, namely, making specific behavioral observations and inferences about the individual patients they treat and know well. It relies on statistical algorithms to do what they do best, namely, combining data optimally to derive reliable and valid diagnostic scales and indexes.

Interrater reliability of SWAP-200 PD scale scores and other diagnostic scales has been above .80 for all scales in all studies to date and has often been above .90 (Marin-Avellan, McGauley, Campbell, & Fonagy, 2004; Westen & Muderrisoglu, 2003, in press). It is noteworthy that high reliability coefficients have been reported by independent investigators unaffiliated with our own laboratory. The reliability coefficients compare favorably with those typically reported for structured interviews that avoid clinical inference and "stick to the facts" (e.g., *DSM-IV* criteria). Additionally, the SWAP diagnostic scales correlate with a wide range of external criterion measures in both adult and adolescent samples including genetic history variables such as psychosis in first- and second-degree relatives, substance abuse in first- and second-degree relatives, developmental history variables such as childhood physical and sexual abuse, life events including psychiatric hospitalizations and suicide attempts, violent criminal behavior, ratings of adaptive functioning, and so on (Marin-Avellan, McGauley, Campbell, & Fonagy, 2005; Shedler & Westen, 2004a; Westen & Muderrisoglu, 2003; Westen & Shedler, 1999a; Westen, Shedler, Durrett, Glass, & Martens, 2003; Westen & Weinberger, 2004).

We describe some illustrative studies.⁵ Westen and Muderrisoglu (2003) interviewed a small sample of outpatients using the Clinical Diagnostic Interview (CDI; Westen, 2002), a systematic interview (2 to 3 hr in length) designed to systematize the personality assessment methods employed by knowledgeable clinicians in real-world practice (Westen, 1997). The CDI does not ask patients to describe their own personality traits but instead elicits narrative descriptions of patients' lives and important relationships. The narrative descriptions allow clinical interviewers to draw inferences about patients' characteristic ways of thinking, feeling, regulating emotion, relating to others, and so on (much as a skilled clinician might do in the first two to four meetings with a new patient).

The primary aims of Westen and Muderrisoglu were (a) to assess interrater reliability of SWAP diagnostic scales as assessed by independent clinicians who either conducted or observed (on videotape) the CDI interview and (b) to assess convergent validity between these independent raters and the treating clinicians whose SWAP scores were based on extensive contact with their patients over time. All clinical assessors were blind to the data provided by the others. The study examined the reliability and validity of the 10 SWAP PD scores and seven additional SWAP diagnostic indexes (see Westen & Shedler, 1999b).

Interrater reliability between independent interviewers averaged greater than .80 for all SWAP diagnostic scales. Median convergent validity coefficients between

⁵The material presented here is adapted from Westen and Weinberger (2004) and updated.

interviewers and treating clinicians were also above .80. Discriminant validity coefficients (i.e., correlations between conceptually unrelated diagnostic scales) were excellent, hovering near 0 for empirically identified diagnostic groupings.⁶ To provide some reference points with which to compare these values, convergent validity between PD diagnoses derived from structured research interviews and diagnoses based on the LEAD (longitudinal evaluation using all available data; Spitzer, 1983) standard have ranged from .00 to .40 in prior studies, and discriminant validity has been notoriously poor (Pilkonis et al., 1995; Pilkonis, Heape, Ruddy, & Serrao, 1991; Widiger & Samuel, 2005). Similarly, a meta-analysis of PD dimensions assessed via self- and informant-report yielded a median correlation of only .36 (Klonsky, Oltmanns, & Turkheimer, 2002).

A second study (Bradley & Westen, 2006) examined convergence between SWAP scores and patient self-report data for borderline and antisocial PD (the two PDs for which self-report and informant-report data tend to converge). Advanced clinical psychology graduate students used the SWAP–200 to describe 54 outpatients after the fifth clinical contact hour. The patients also completed the Personality Assessment Inventory (PAI; Morey, 1991). Convergent validity was high, with SWAP antisocial and borderline PD scores differentially predicting antisocial and borderline scores on the PAI, and discriminant validity coefficients were desirably low.

A study from a research group other than our own reported comparable findings (Marin-Avellan et al., 2005). Marin-Avellan et al. applied the SWAP–200 to audiotaped Adult Attachment Interviews (AAI; Main, Kaplan, & Cassidy, 1985) plus chart records for a sample of inpatients at a maximum security forensic hospital. Interrater reliability between independent assessors was high for all SWAP–200 PD scales, with a median interrater correlation of $r = .91$.

The SWAP antisocial PD score differentiated patients who had committed violent versus nonviolent offenses, whereas SCID–II scores did not. The SWAP–200 also proved superior to the SCID–II in predicting aggressive ward behavior, independently assessed by ward nurses (blind to other data) using a 49-item interpersonal circumplex rating scale. SWAP antisocial PD scores correlated significantly with dominance behavior and coercive behavior on the ward and correlated negatively with submissive behavior and compliant behavior on the ward. In contrast, the SCID–II predicted only dominance behavior. The findings demonstrate incremental validity of the SWAP–200 relative to the gold standard structured research interview. Overall, the research indicates that experienced clinicians can make reliable and valid observations and inferences about personality dynamics given a suitable technology for harnessing their judgments.

TOWARD *DSM–V*: AN IMPROVED CLASSIFICATION OF PERSONALITY DISORDERS⁷

It is an empirical question whether the *DSM–IV* includes the optimal diagnostic categories and criteria. It is also an empirical question whether a diagnostic system based on personality types or syndromes is consistent with the available data (as opposed to being a mere convenience that facilitates clinical communication). To address these questions, Westen and Shedler (1999b) examined SWAP–200 personality descriptions from a national sample of patients ($n = 496$) diagnosed with Axis II disorders provided by the treating psychologist or psychiatrist. We used the technique of Q-factor analysis (or simply Q-analysis) to answer the following questions:

1. Are there clear, empirically identifiable diagnostic groupings among PD patients treated in the community? That is, are there groupings of patients who share common psychological features that distinguish them from other patients?
2. Do the current *DSM–IV* diagnostic categories adequately “fit” the data? That is, are there empirically identifiable personality syndromes that are not included in *DSM–IV* or vice-versa?
3. What are the most defining psychological features (diagnostic criteria) for each personality syndrome?

Q-analysis is computationally equivalent to the familiar technique of factor analysis. The difference is that factor analysis identifies groupings of similar *variables* (i.e., columns in a data matrix) that are assumed to be markers of a common underlying factor. In contrast, Q-analysis identifies groupings of similar *people* (i.e., cases or rows in a data matrix) who are assumed to represent a common diagnostic syndrome or type. The former approach is *variable centered*, and the latter is *person centered* (for a description of the computational methods, see Westen & Shedler, 1999b). Q-analysis has been used by biologists conducting taxonomic research to aid in classifying species and has been used successfully in research on normal personality (Block, 1971).⁸

⁷The material presented here is adapted from Westen and Shedler (1999b) and updated.

⁸A discussion of person-centered versus variable-centered assessment is beyond the scope of this article and warrants a paper in its own right. We believe the distinction underlies much misunderstanding between clinicians and researchers because clinicians tend to think in person-centered terms and researchers tend to think in variable-centered terms. The choice of a person- or variable-centered approach, which may profoundly affect how we think about psychological issues, is often not even recognized as a choice. Instead, one or the other approach is accepted by convention and without consideration of what is at stake. (The fact that SPSS is designed to manipulate variables rather than cases may have shaped academic psychology in ways we can barely fathom). It is not that one approach is “right” and one is “wrong” but rather that they serve different purposes and draw our

⁶All coefficients are Pearson’s r .

The Q-analysis analysis demonstrated that there are empirically distinguishable personality syndromes among patients treated in the community and that a syndromal (or person-centered) approach is consistent with the data. Note that the personality syndromes are best understood dimensionally and not as mutually exclusive categories (see the SWAP Dimensional Diagnosis section previously). This is an important clarification because some researchers mistakenly conflate dimensional with trait (variable-centered) models and conflate categorical approaches with syndromal (person-centered) models. In fact, these issues are separate and independent (Westen, Gabbard, & Blagov, 2006). The dimensional–categorical distinction refers to whether people are assumed to fall into discrete categories or to vary along a continuum; the syndromal–trait distinction refers to whether the unit of diagnosis is a constellation of interrelated personality characteristics or separate characteristics.

The Q-analysis identified 11 conceptually coherent diagnostic groupings or personality syndromes, many of which resembled *DSM-IV* diagnostic categories and some of which did not. We created a prototype personality description for each empirically identified diagnostic syndrome by listing the SWAP items in descending order by Q-factor score. The SWAP–200 items with the highest Q-factor scores indicate the central or defining psychological features of each diagnostic group (i.e., the diagnostic criteria). This represents a purely empirical approach to identifying optimal diagnostic categories and criteria.

We use the examples of paranoid and depressive PD to illustrate this approach to identifying PD syndromes and criteria (for descriptions of all 11 empirically identified diagnostic groupings, see Shedler & Westen, 2004b; Westen & Shedler, 1999b). Table 1 lists the SWAP–200 items most defining of patients in the paranoid personality diagnostic grouping along with their associated factor scores (indicating their diagnostic importance). A number of findings are noteworthy. First, the empirical identification of this diagnostic grouping validates the inclusion of paranoid PD as a diagnostic category in *DSM-IV*. Second, the items are clinically richer than the *DSM-IV* diagnostic criteria, addressing inner experience as well as behavior. Third, the description differs in important ways from the description provided by *DSM-IV* and offers crucial insights into the meaning and function of paranoid symptoms that are absent from *DSM*.

For example, the prototype description (Table 1) emphasizes paranoid patients' cognitive confusion in ways that *DSM-IV* does not (e.g., tends to become irrational when strong emotions are stirred up, has difficulty making sense

attention to different matters. Good assessment systems are like good maps, in that they must accurately depict the territory. However, sometimes one wants a roadmap, sometimes a map of elevations, and sometimes a political map. The motorist trying to navigate the interstate will have little interest in a map of elevations irrespective of the number of studies documenting its reliability and validity.

TABLE 1
Prototype Description of Patients in the
Paranoid Personality Disorder Diagnostic
Category

<i>Item</i>	<i>Score</i>
Tends to hold grudges; may dwell on insults or slights for long periods.	3.61
Tends to feel misunderstood, mistreated, or victimized.	3.23
Is quick to assume that others wish to harm or take advantage of him/her; tends to perceive malevolent intentions in others' words and actions.	3.08
Tends to express intense and inappropriate anger, out of proportion to the situation at hand.	2.77
Tends to be critical of others.	2.59
Tends to get into power struggles.	2.43
Tends to be angry or hostile (whether consciously or unconsciously).	2.40
Tends to see certain others as "all bad," and loses the capacity to perceive any positive qualities the person may have.	2.38
Tends to be self-righteous or moralistic.	2.25
Tends to react to criticism with feelings of rage or humiliation.	2.19
Tends to blame others for own failures or shortcomings; tends to believe his/her problems are caused by external factors.	2.15
Tends to be oppositional, contrary, or quick to disagree.	2.08
Tends to see own unacceptable feelings or impulses in other people instead of in him/herself.	2.08
Tends to become irrational when strong emotions are stirred up; may show a noticeable decline from customary level of functioning.	1.94
Tends to "catastrophize"; is prone to see problems as disastrous, unsolvable, etc.	1.82
Tends to elicit dislike or animosity in others.	1.78
Emotions tend to spiral out of control, leading to extremes of anxiety, sadness, rage, excitement, etc.	1.73
Has difficulty making sense of other people's behavior; often misunderstands, misinterprets, or is confused by others' actions and reactions.	1.53
Tends to be controlling.	1.49
Tends to elicit extreme reactions or stir up strong feelings in others.	1.44
Tends to avoid confiding in others for fear of betrayal; expects things s/he says or does will be used against him/her.	1.41
Reasoning processes or perceptual experiences seem odd and idiosyncratic (e.g., may make seemingly arbitrary inferences; may see hidden messages or special meanings in ordinary events).	1.39
Perception of reality can become grossly impaired under stress (e.g., may become delusional).	1.37

Note. From "Revising and Assessing Axis II: II. Toward an Empirically Based and Clinically Useful Classification of Personality Disorders," by D. Westen and J. Shedler, 1999, *American Journal of Psychiatry*, 156, p. 278. Copyright 1999 by the American Psychiatric Association. Reprinted with permission.

of other people's behavior, reasoning processes or perceptual experiences seem odd and idiosyncratic). It also emphasizes paranoid patients' anger and hostility in ways that *DSM* does not (e.g., tends to hold grudges, tends to be angry or hostile, tends to express intense and inappropriate anger), as well as their tendency to rely on projection as a defense (tends to see own unacceptable feelings or impulses in other people instead of in himself/herself). The findings are consistent with long-held clinical theories that recognize projection as

a central dynamic in paranoid patients (stated differently, the paranoid patient perceives the world as hostile because the patient sees his or her own hostility wherever he or she looks). The personality description has clear implications for treatment, unlike the description in *DSM–IV*. It makes clear that a clinician treating a patient with paranoid PD will need to assist the patient with reality testing, for example, by examining his reasoning processes and helping him to consider alternate constructions and interpretations of events. It also makes clear that the clinician will be dealing with intense anger and hostility, and that any successful treatment will have to address the patient’s anger and help him to find more adaptive ways of expressing and regulating it.

The findings cannot be explained away as mere artifacts of clinicians’ theoretical beliefs or expectations. They emerged repeatedly when we stratified the sample by the theoretical orientation of the reporting clinicians, and the personality characteristics described previously were ranked just as highly by cognitive-behavioral therapists as by psychodynamic therapists. The SWAP–200 provides a common language for all clinicians, and the PD prototypes reflect only those personality traits that clinicians of all orientations observe consistently and reliably.

Table 2 lists the SWAP–200 items most defining of another personality syndrome, one absent from *DSM–IV*, which we have labeled *depressive* (or *dysphoric*) personality. Despite its omission from *DSM–IV*, our data indicate that it is the most prevalent personality syndrome seen in clinical practice. Its absence from *DSM–IV* appears to be a significant omission. Note that the SWAP description encompasses the *multiple domains of functioning* described in *DSM–IV* as defining of a PD including cognition (e.g., tends to blame self, tends to be self-critical), affectivity (e.g., tends to feel unhappy, depressed, despondent; tends to feel ashamed embarrassed), interpersonal relations (tends to fear she or he will be rejected or abandoned, tends to be overly needy or dependent), and impulse regulation (e.g., has difficulty acknowledging or expressing anger). The syndrome appears to have its origin in late childhood or early adolescence (Westen et al., 2005) and appears stable and enduring over time. In short, it appears to be a PD by every definition of the term (cf. Huprich, 2003, 2005; Huprich & Frisch, 2004; McDermut, Zimmerman, & Chelminski, 2003).

Analysis of SWAP–II data from our most recent subject samples ($N = 1,201$ adult patients studied with the SWAP–II, and $N = 950$ adolescent patients with the SWAP–II–A) revealed a hierarchical structure of PD syndromes as illustrated in Figure 2. At the superordinate level are three broad diagnostic groupings (Q-factors) that can be described as internalizing, externalizing, and borderline. The results map onto the internalizing and externalizing spectra identified in research on Axis I syndromes (Krueger, 2002) and may provide a basis for an integrated understanding of Axis I and Axis II pathology. The borderline personality constellation contains elements of both internalizing and externalizing

TABLE 2
Prototype Description of Patients in the
Depressive (Dysphoric) Personality
Disorder Diagnostic Category

<i>Item</i>	<i>Score</i>
Tends to feel s/he is inadequate, inferior, or a failure.	3.63
Tends to feel unhappy, depressed, or despondent.	3.11
Tends to feel ashamed or embarrassed.	2.76
Tends to blame self or feel responsible for bad things that happen.	2.71
Tends to feel guilty.	2.67
Tends to fear s/he will be rejected or abandoned by those who are emotionally significant.	2.66
Tends to feel helpless, powerless, or at the mercy of forces outside his/her control.	2.52
Tends to be overly needy or dependent; requires excessive reassurance or approval.	2.30
Tends to be ingratiating or submissive (e.g., may consent to things s/he does not agree with or does not want to do, in the hope of getting support or approval).	2.12
Tends to be passive and unassertive.	2.12
Tends to be self-critical; sets unrealistically high standards for self and is intolerant of own human defects.	2.02
Tends to feel like an outcast or outsider; feels as if s/he does not truly belong.	1.94
Tends to be anxious.	1.91
Tends feel listless, fatigued, or lacking in energy.	1.79
Tends to feel empty or bored.	1.77
Appears to want to “punish” self; creates situations that lead to unhappiness, or actively avoids opportunities for pleasure and gratification.	1.71
Appears to find little or no pleasure, satisfaction, or enjoyment in life’s activities.	1.71
Tends to be insufficiently concerned with meeting own needs; appears not to feel entitled to get or ask for things s/he deserves.	1.70
Is unable to soothe or comfort self when distressed; requires involvement of another person to help regulate affect.	1.60
Lacks a stable image of who s/he is or would like to become (e.g., attitudes, values, goals, and feelings about self may be unstable and changing).	1.55
Tends to feel life has no meaning.	1.49
Tends to avoid social situations because of fear of embarrassment or humiliation.	1.43
Has difficulty acknowledging or expressing anger.	1.34

Note. From “Revising and Assessing Axis II: II. Toward an Empirically Based and Clinically Useful Classification of Personality Disorders,” by D. Westen and J. Shedler, 1999, *American Journal of Psychiatry*, 156, 278. Copyright 1999 by American Psychiatric Association. Reprinted with permission.

pathology and is characterized by emotional instability that is not observed in stable internalizers or externalizers. The hierarchical structure of PD syndromes will be described in greater detail in future publications.

DIMENSIONAL DIAGNOSIS: THE “PROTOTYPE MATCHING” APPROACH

For research purposes, for situations in which maximum psychometric precision is required (as in forensic assessment),

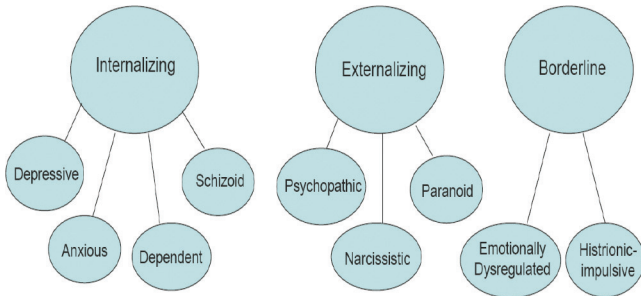


FIGURE 2. Hierarchical structure of personality syndromes.

or where there are challenging diagnostic dilemmas, clinicians can describe patients using the SWAP–200 and obtain dimensional diagnosis scores such as those graphed in Figure 1. (Investigators will soon be able to enter SWAP data and receive comprehensive diagnostic reports via the Internet; for information, visit www.SWAPassessment.org). Where routine use of the SWAP–200 would be impractical, we advocate a prototype matching approach to personality diagnosis. We have proposed this approach for *DSM–V*.

Figure 3 illustrates the prototype matching approach. The figure shows the prototype description for one personality syndrome, identified empirically through Q-analysis, labeled *antisocial-psychopathic personality* (Westen & Shedler, 1999b). The description is made up of the SWAP–200 statements that are most empirically defining of the syndrome. The SWAP items are reproduced essentially verbatim but have been arranged in paragraph (rather than list) form.

The clinician’s task is to consider the prototype description as a whole—that is, as a configuration or gestalt—and to rate the overall similarity or match between the prototype and the patient at hand. The resulting diagnosis is dimensional (a 1–5 rating), but the scale can be dichotomized for convenience when a present/absent classification is desired to facilitate clinical communication (with a rating ≥ 4 indi-

ating/“caseness”). Thus, the approach offers the advantages of dimensional diagnosis while maintaining backward compatibility with the categorical approach of *DSM–IV*.

Our research indicates that the prototype matching method has substantial advantages over the current *DSM–IV* approach to personality diagnosis. In a series of studies of Cluster B disorders (antisocial, borderline, histrionic, and narcissistic), we compared prototype matching and *DSM–IV* diagnosis with respect to validity, diagnostic comorbidity, and clinical utility (Westen, Shedler, & Bradley, 2006). Clinicians diagnosed patients using the prototype matching method (as illustrated in Figure 3) and using the *DSM–IV* diagnostic system. We compared the prototype matching method to both categorical *DSM–IV* diagnoses and “dimensionalized” *DSM–IV* diagnoses obtained by summing the number of criteria met per disorder (a method commonly used in PD research).

The prototype matching method substantially reduced diagnostic comorbidity relative to both *DSM–IV* diagnostic methods. For example, the median correlation between the four Cluster B disorders was $r = .47$ for dimensionalized *DSM–IV* diagnoses and $r = .14$ for prototype diagnoses based on the empirically identified diagnostic groupings (as described in the preceding section). At the same time, the prototype diagnoses had higher validity, yielding higher correlations with ratings of adaptive functioning and with developmental history variables known to be associated with antisocial PD and borderline PD. The advantages of the prototype matching approach were not only statistically significant but also clinically meaningful: Prototype diagnoses predicted treatment outcomes better than either categorical or dimensionalized *DSM–IV* diagnoses, for both psychotherapy and antidepressant medication.

Finally, we examined clinical utility by asking clinicians to compare the prototype matching method to the *DSM–IV* diagnostic system with respect to ease of use, usefulness for communicating with other clinicians, ability to capture important information about the patient, and so on. The clinicians strongly preferred prototype diagnosis to *DSM–IV* diagnosis on every dimension assessed despite the fact that they had no prior experience with either the prototype matching method or the empirically derived prototype personality descriptions. Spitzer, First, et al. (2006) also found that clinicians strongly prefer prototype matching to both *DSM–IV* diagnosis and dimensional trait models such as the Five Factor Model, and rate the approach as more clinically relevant and useful.

Patients who match this prototype tend to be deceitful, to lie and mislead people. They take advantage of others, have minimal investment in moral values, and appear to experience no remorse for harm or injury caused to others. They tend to manipulate others’ emotions to get what they want; to be unconcerned with the consequences of their actions, appearing to feel immune or invulnerable; and to show reckless disregard for the rights, property, or safety of others. They have little empathy, and seem unable to understand or respond to others’ needs and feelings unless they coincide with their own. Individuals who match this prototype tend to act impulsively, without regard for consequences; to be unreliable and irresponsible (e.g., failing to meet work obligations or honor financial commitments); to engage in unlawful or criminal behavior; and to abuse alcohol. They tend to be angry or hostile; to get into power struggles; and to gain pleasure or satisfaction by being sadistic or aggressive toward others. Patients who match this prototype tend to blame others for their own failures or shortcomings, and to believe their problems are caused by external factors. They have little psychological insight into their own motives, behavior, etc. They may repeatedly convince others of their commitment to change but then revert to previous maladaptive behavior, often convincing others that “this time is really different.”

Please form an overall impression of the type of person described, then rate the extent to which your patient matches or resembles this prototype.

5 very good match (patient exemplifies this disorder; prototypical case)	Diagnosis
4 good match (patient has this disorder; diagnosis applies)	
3 moderate match (patient has significant features of this disorder)	Features
2 slight match (patient has minor features of this disorder)	
1 no match (description does not apply)	

FIGURE 3. Antisocial-psychopathic personality disorder prototype.

CONCLUSION: INTEGRATING SCIENCE AND PRACTICE

A clinically useful diagnostic system should encompass the spectrum of personality pathology seen in clinical practice and have meaningful implications for treatment.

An empirically sound diagnostic system should facilitate reliable and valid diagnoses: Independent clinicians should be able to arrive at the same diagnosis, the diagnoses should be relatively distinct from one another, and each diagnosis should be associated with unique and theoretically meaningful correlates, antecedents, and sequelae (Livesley & Jackson, 1992; Millon, 1991; Robins & Guze, 1970).

One obstacle to achieving this ideal has been an unfortunate schism in the mental health professions between science and practice. Too often, research has been conducted in isolation from the crucial data of clinical observation. The results often strike clinicians as naïve and of dubious clinical relevance. Ultimately, the most empirically elegant diagnostic system will have little impact if clinicians do not find it helpful for understanding their patients (First et al., 2004; Shedler & Westen, 2005). On the other hand, clinical theory has too often developed without sufficient regard for questions of falsifiability and empirical credibility. The results have often struck researchers as scientifically naïve.

The SWAP represents an effort to bridge the schism between science and practice by quantifying clinical observation and expertise, making clinical constructs accessible to empirical study. It relies on clinicians to do what they do best, namely, making observations and inferences about the individual patients they know and treat. It relies on quantitative methods to do what they do best, namely, aggregating observations to reveal relationships and commonalities and combining data to yield optimal predictions (cf. Sawyer, 1966). It provides a language for clinical case description that is at once psychometrically sound and clinically rich enough to describe the complexities of real patients. There remains a sizeable schism between science and practice. Perhaps the SWAP will provide a language all parties can speak.

ACKNOWLEDGMENT

This article was adapted with permission from material previously published in J. Shedler & D. Westen, D., “Personality diagnosis with the Shedler–Westen Assessment Procedure (SWAP): Bridging the gulf between science and practice,” in *Psychodynamic Diagnostic Manual (PDM)*, by Alliance Task Force (Ed.), 2006, Silver Spring, MD: Alliance of Psychoanalytic Organizations. Copyright 2006 by Alliance Task Force.

REFERENCES

- American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.
- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- Betan, E., Heim, A. K., Conklin, C. Z., & Westen, D. (2005). Countertransference phenomena and personality pathology in clinical practice: an empirical investigation. *American Journal of Psychiatry*, *162*, 890–898.
- Blais, M., & Norman, D. (1997). A psychometric evaluation of the DSM–IV personality disorder criteria. *Journal of Personality Disorders*, *11*, 168–176.
- Blashfield, R. (1985). Exemplar prototypes of personality disorder diagnoses. *Comprehensive Psychiatry*, *26*, 11–21.
- Block, J. (1971). *Lives through time*. Berkeley, CA: Bancroft.
- Block, J. (1978). *The Q-sort method in personality assessment and psychiatric research*. Palo Alto, CA: Consulting Psychologists Press. (Original work published 1961)
- Bradley, R., & Westen, D. (2006). *Validity of SWAP–200 personality diagnosis in an outpatient sample*. Unpublished manuscript, Emory University.
- Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A. M., & Kaemmer, B. (1989). *MMPI–2: Manual for administration and scoring*. Minneapolis: University of Minnesota.
- Cantor, N., & Genero, N. (1986). Psychiatric diagnosis and natural categorization: A close analogy. In T. Millon & G. L. Klerman (Eds.), *Contemporary directions in psychopathology: Toward the DSM–IV* (pp. 233–256). New York: Guilford.
- Clark, L. (1992). Resolving taxonomic issues in personality disorders: the value of larger scale analyses of symptom data. *Journal of Personality Disorders*, *6*, 360–376.
- Clarkin, J. F., Levy, K. N., Lenzenweger, M. F., & Kernberg, O. F. (2007). *The Personality Disorders Institute/Borderline Personality Disorder Research Foundation randomized control trial for borderline personality disorder: Treatment outcome*. Manuscript submitted for publication.
- Cousineau, T. M., & Shedler, J. (2006). Predicting physical health: Implicit mental health measures versus self report scales. *Journal of Nervous and Mental Disease*, *194*, 427–432.
- Davis, R., Blashfield, R., & McElroy, R. (1993). Weighting criteria in the diagnosis of a personality disorder: a demonstration. *Journal of Abnormal Psychology*, *102*, 319–322.
- First, M. B., Pincus, H. A., Levine, J. B., Williams, J. B. W., Uston, B., & Peel, R. (2004). Clinical utility as a criterion for revising psychiatric diagnoses. *American Journal of Psychiatry*, *161*, 946–954.
- First, M., Spitzer, R., Gibbon, M., Williams, J., Davies, J. B., Howes, M., et al. (1995). The Structured Clinical Interview for DSM–III–R Personality Disorders (SCID–II): Part II. Multi-site test-retest reliability study. *Journal of Personality Disorders*, *9*, 92–104.
- Grilo, C. M., Sanislow, C. A., & McGlashan, T. H. (2002). Co-occurrence of DSM–IV personality disorders with borderline personality disorder. *Journal of Nervous and Mental Disease*, *190*, 552–553.
- Grove, W. M., & Tellegen, A. (1991). Problems in the classification of personality disorders. *Journal of Personality Disorders*, *5*, 31–41.
- Harkness, A. (1992). Fundamental topics in the personality disorders: Candidate trait dimensions from lower regions of the hierarchy. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, *4*, 251–259.
- Hathaway, S. R., & McKinley, J. C. (1943). *The Minnesota Multiphasic Personality Inventory*. Minneapolis: University of Minnesota Press.
- Huprich, S. K. (2003). Depressive personality and its relationship to depressed mood, interpersonal loss, negative parental perceptions, and perfectionism. *Journal of Nervous and Mental Disease*, *191*, 1–7.
- Huprich, S. K. (2005). Differentiating avoidant and depressive personalities. *Journal of Personality Disorders*, *19*, 659–673.
- Huprich, S. K., & Frisch, M. B. (2004). Depressive personality and its relationship to quality of life, hope, and optimism. *Journal of Personality Assessment*, *83*, 22–28.
- Jackson, D. N., & Livesley, W. J. (1995). Possible contributions from personality assessment to the classification of personality disorders. In W. J. Livesley (Ed.), *The DSM–IV personality disorders* (pp. 459–481). New York: Guilford.
- John, O. (1990). The “big five” factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In L. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 66–100). New York: Guilford.

- Kernberg, O. (1975). *Borderline conditions and pathological narcissism*. New York: Aronson.
- Kernberg, O. (1984). *Severe personality disorders*. New Haven, CT: Yale University Press.
- Kim, N. S., & Ahn, W. (2002). Clinical psychologists' theory-based representations of mental disorders predict their diagnostic reasoning and memory. *Journal of Experimental Psychology*, *131*, 451–476.
- Klonsky, E. D., Oltmanns, T. F., & Turkheimer, E. (2002). Informant-reports of personality disorder: Relation to self-reports and future research directions. *Clinical Psychology: Science & Practice*, *9*, 300–311.
- Kohut, H. (1971). *The analysis of the self*. New York: International Universities Press.
- Krueger, R. F. (2002). The structure of common mental disorders. *Archives of General Psychiatry*, *59*, 570–571.
- Levy, K. N., Meehan, K. B., Kelly, K. M., Reynoso, J. S., Weber, M., Clarkin, J. F., et al. (2006). Change in attachment patterns and reflective function in a randomized control trial of transference-focused psychotherapy for borderline personality disorder. *Journal of Consulting and Clinical Psychology*, *74*, 1027–1040.
- Linehan, M. M. (1993). *Cognitive-behavioral treatment of borderline personality disorder*. New York: Guilford.
- Lingiardi, V., Shedler, J., & Gazillo, F. (2006). Assessing personality change in psychotherapy with the SWAP–200: A case study. *Journal of Personality Assessment*, *86*, 23–32.
- Livesley, W. J. (Ed.). (1995). *The DSM–IV personality disorders*. New York: Guilford.
- Livesley, W. J., & Jackson, D. N. (1992). Guidelines for developing, evaluating, and revising the classification of personality disorders. *Journal of Nervous and Mental Disease*, *180*, 609–618.
- Main, M., Kaplan, N., & Cassidy, J. (1985). Security in infancy, childhood, and adulthood: A move to the level of representation. *Monographs of the Society for Research in Child Development*, *50*(1–2, Serial No. 209).
- Marin-Avellan, L., McGauley, G., Campbell, C., & Fonagy, P. (2005). Using the SWAP–200 in a personality-disordered forensic population: Is it valid, reliable and useful? *Journal of Criminal Behaviour and Mental Health*, *15*, 28–45.
- McCrae, R., & Costa, P. (1990). *Personality in adulthood*. New York: Guilford.
- McDermut, W., Zimmerman, M., & Chelminski, I. (2003). The construct validity of depressive personality disorder. *Journal of Abnormal Psychology*, *112*, 49–60.
- McWilliams, N. (1994). *Psychoanalytic diagnosis: Understanding personality structure in the clinical process*. New York: Guilford.
- Millon, T. (1990). *Toward a new psychology*. New York: Wiley.
- Millon, T. (1991). Classification in psychopathology: Rationale, alternatives and standards. *Journal of Abnormal Psychology*, *100*, 245–261.
- Millon, T., & Davis, R. D. (1997). The place of assessment in clinical science. In T. Millon (Ed.), *The Millon inventories: Clinical and personality assessment* (pp. 3–20). New York: Guilford.
- Morey, L. C. (1988). Personality disorders in *DSM–III* and *DSM–III–R*: Convergence, coverage, and internal consistency. *American Journal of Psychiatry*, *145*, 573–577.
- Morey, L. C. (1991). *The Personality Assessment Inventory: Professional manual*. Odessa, FL: Psychological Assessment Resources.
- Oldham, J., Skodol, A., Kellman, H. D., Hyler, S., Rosnick, L., & Davies, M. (1992). Diagnosis of *DSM–III–R* personality disorders by two semistructured interviews: Patterns of comorbidity. *American Journal of Psychiatry*, *149*, 213–220.
- Perry, J. C. (1992). Problems and considerations in the valid assessment of personality disorders. *American Journal of Psychiatry*, *149*, 1645–1653.
- Perry, J. C., & Cooper, S. H. (1987). Empirical studies of psychological defense mechanisms. In J. Cavenar & R. Michels (Eds.) *Psychiatry*. Philadelphia: Lippincott.
- Pilkonis, P. A., Heape, C. L., Proietti, J. M., Clark, S. W., McDavid, J. D., & Pitts, T. E. (1995). The reliability and validity of two structured diagnostic interviews for personality disorders. *Archives of General Psychiatry*, *52*, 1025–1033.
- Pilkonis, P. A., Heape, C. L., Ruddy, J., & Serrao, P. (1991). Validity in the diagnosis of personality disorders: The use of the LEAD standard. *Psychological Assessment*, *31*, 46–54.
- Robins, E., & Guze, S. (1970). The establishment of diagnostic validity in psychiatric illness: Its application to schizophrenia. *American Journal of Psychiatry*, *126*, 983–987.
- Sawyer, J. (1966). Measurement and prediction, clinical and statistical. *Psychological Bulletin*, *66*, 178–200.
- Shedler, J., Mayman, M., & Manis, M. (1993). The illusion of mental health. *American Psychologist*, *48*, 1117–1131.
- Shedler, J., & Westen, D. (1998). Refining the measurement of Axis II: A Q-sort procedure for assessing personality pathology. *Assessment*, *5*, 333–353.
- Shedler, J., & Westen, D. (2004a). Dimensions of personality pathology: An alternative to the five factor model. *American Journal of Psychiatry*, *161*, 1743–1754.
- Shedler, J., & Westen, D. (2004b). Refining *DSM–IV* personality disorder diagnosis: Integrating science and practice. *American Journal of Psychiatry*, *161*, 1350–1365.
- Shedler, J., & Westen, D. (2005). A simplistic view of the Five-factor model: Drs. Shedler and Westen reply. *American Journal of Psychiatry*, *162*, 1551.
- Shedler, J., & Westen, D. (2006). Personality diagnosis with the Shedler–Westen Assessment Procedure (SWAP): Bridging the gulf between science and practice. In Alliance Task Force (Ed.), *Psychodynamic diagnostic manual (PDM)* (pp. 573–613). Silver Spring, MD: Alliance of Psychoanalytic Organizations.
- Skodol, A., Oldham, J., Rosnick, L., Kellman, D., & Hyler, S. (1991). Diagnosis of *DSM–III–R* personality disorders: A comparison of two structured interviews. *International Journal of Methods in Psychiatric Research*, *1*, 13–26.
- Spitzer, R. L. (1983). Psychiatric diagnosis: Are clinicians still necessary? *Comprehensive Psychiatry*, *24*, 399–411.
- Spitzer, R. L., First, M. B., Shedler, J., Westen, D., & Skodol, A. (2006). *Clinical utility of five dimensional systems for personality diagnosis: A "consumer preference" study*. Manuscript submitted for publication.
- Vaillant, G. (Ed.). (1992). *Ego mechanisms of defense: A guide for clinicians and researchers*. Washington, DC: American Psychiatric Press.
- Watson, D., & Sinha, B. K. (1998). Comorbidity of *DSM–IV* personality disorders in a nonclinical sample. *Journal of Clinical Psychology*, *54*, 773–780.
- Westen, D. (1991). Social cognition and object relations. *Psychology Bulletin*, *109*, 429–455.
- Westen, D. (1997). Divergences between clinical and research methods for assessing personality disorders: Implications for research and the evolution of Axis II. *American Journal of Psychiatry*, *154*, 895–903.
- Westen, D. (2002). *Clinical Diagnostic Interview*. Unpublished manual, Emory University. Retrieved June 4, 2007, from www.psychsystems.net/lab.
- Westen, D., & Arkowitz-Western, L. (1998). Limitations of Axis II in diagnosing personality pathology in clinical practice. *American Journal of Psychiatry*, *155*, 1767–1771.
- Westen, D., Dutra, L., & Shedler, J. (2005). Assessing adolescent personality pathology. *British Journal of Psychiatry*, *186*, 227–238.
- Westen, D., Gabbard, G., & Blagov, P. (2006). Back to the future: Personality structure as a context for psychopathology. In R. F. Krueger & J. L. Tackett (Eds.), *Personality and psychopathology* (pp. 335–384). New York: Guilford.
- Westen, D., Heim, A. K., Morrison, K., Patterson, M., & Campbell, L. (2002). Classifying and diagnosing psychopathology: A prototype matching approach. In L. E. Beutler & M. Malik (Eds.), *Rethinking the DSM*:

- A psychological perspective* (pp. 221–250). Washington, DC: American Psychological Association Press.
- Westen, D., Lohr, N., Silk, K., Gold, L., & Kerber, K. (1990). Object relations and social cognition in borderlines, major depressives, and normals: A TAT analysis. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, *2*, 355–364.
- Westen, D., & Muderrisoglu, S. (2003). Reliability and validity of personality disorder assessment using a systematic clinical interview: Evaluating an alternative to structured interviews. *Journal of Personality Disorders*, *17*, 350–368.
- Westen, D., & Muderrisoglu, S. (in press). Clinical assessment of pathological personality traits. *American Journal of Psychiatry*.
- Westen, D., Muderrisoglu, S., Fowler, C., Shedler, J., & Koren, D. (1997). Affect regulation and affective experience: Individual differences, group differences, and measurement using a Q-sort procedure. *Journal of Consulting and Clinical Psychology*, *65*, 429–439.
- Westen, D., & Shedler, J. (1999a). Revising and assessing Axis II: I. Developing a clinically and empirically valid assessment method. *American Journal of Psychiatry*, *156*, 258–272.
- Westen, D., & Shedler, J. (1999b). Revising and assessing Axis II: II. Toward an empirically based and clinically useful classification of personality disorders. *American Journal of Psychiatry*, *156*, 273–285.
- Westen, D., & Shedler, J. (2000). A prototype matching approach to diagnosing personality disorders toward DSM–V. *Journal of Personality Disorders*, *14*, 109–126.
- Westen, D., Shedler, J., & Bradley, R. (2006). A prototype approach to personality diagnosis. *American Journal of Psychiatry*, *163*, 846–856.
- Westen, D., Shedler, J., Durrett, C., Glass, S., & Martens, A. (2003). Personality diagnosis in adolescence: DSM–IV Axis II diagnoses and an empirically derived alternative. *American Journal of Psychiatry*, *160*, 952–966.
- Westen, D. W., Waller, N., Shedler, J., Blagov, P., & Bradley, R. (2006). The structure of personality traits using the SWAP–II: Data from a large normative sample. Under review.
- Westen, D., & Weinberger, J. (2004). When clinical description becomes statistical prediction. *American Psychologist*, *59*, 595–613.
- Widiger, T. A. (1993). The DSM–III–R categorical personality disorder diagnoses: A critique and an alternative. *Psychological Inquiry*, *4*, 75–90.
- Widiger, T. A., & Corbitt, E. M. (1995). Antisocial personality disorder. In J. W. Livesley (Ed.), *The DSM–IV personality disorders: Diagnosis and treatment of mental disorders* (pp. 103–126). New York: Guilford.
- Widiger, T., & Frances, A. (1985). The DSM–III personality disorders: Perspectives from psychology. *Archives of General Psychiatry*, *42*, 615–623.
- Widiger, T. A., & Samuel, D. B. (2005). Evidence-based assessment of personality disorders. *Psychological Assessment*, *17*, 278–287.
- Widiger, T. A., & Simonsen, E. S. (2005). Alternative dimensional models of personality disorders: Finding common ground. *Journal of Personality Disorders*, *19*, 110–130.
- Zimmerman, M. (1994). Diagnosing personality disorders: A review of issues and research methods. *Archives of General Psychiatry*, *51*, 225–245.

Jonathan Shedler
 Department of Psychiatry
 University of Colorado Health Sciences Center
 4455 E. 12th Avenue
 Denver, CO 80220
 Email: jonathan@shedler.com

Received March 28, 2006
 Revised February 20, 2007