

## BRIEF REPORTS

# Relationship Between Clinician Assessment and Self-Assessment of Personality Disorders Using the SWAP-200 and PAI

Rebekah Bradley  
Emory University

Mark Hilsenroth  
Adelphi University

Clifford Guarnaccia and Drew Westen  
Emory University

The relation between self- and peer-informant reports of personality using psychometric instruments has been the focus of considerable research. The quantified judgments of clinically experienced observers such as treating clinicians have also been studied. The focus of the present article is on the measurement of 3 personality disorders (borderline, antisocial, and obsessive–compulsive) using the Shedler–Westen Assessment Procedure (SWAP-200), an instrument designed to quantify personality ratings made by clinically experienced informants, and the self-report Personality Assessment Inventory (PAI). SWAP-200 personality disorder scales showed small to medium correlations with borderline and antisocial personality disorder scales from the PAI. As predicted, SWAP-200 obsessive–compulsive personality disorder correlated negatively with these scales, suggesting discriminant validity.

*Keywords:* SWAP-200, personality disorders, systematic clinical interview, Q-sort, clinical prediction

The development of structured interviews and questionnaires to assess personality disorders (PDs) has led to an enormous growth of research on PDs over the past 25 years. This research has yielded multiple methods and instruments for assessing personality pathology. The methods vary across informants, including self-, peer, and clinician report. Research findings indicate that data gathered across rater type show significant but generally small to moderate correlations, a finding similar to findings from research in other domains of psychopathology (Clifton, Turkheimer, & Oltmanns, 2005; Klein, 2003; Klonsky, Oltmanns, & Turkheimer, 2002; Oltmanns, Melley, & Turkheimer, 2002; Thomas, Turkheimer, & Oltmanns, 2003). The magnitude of those correlations may depend on the disorder, with some disorders (e.g., borderline; BPD) showing stronger convergence between self- and informant-reports and others (e.g., narcissistic PD) showing virtually none (e.g., see Thomas et al., 2003).

Turkheimer, Oltmanns, and their colleagues have suggested that self-reports may be more accurate measures of internalizing pa-

thology (which individuals may not share with acquaintances), whereas peer reports may be more accurate measures of externalizing pathology (which individuals may not themselves recognize; Turkheimer, personal communication, March 2001). Although an increasing body of research has focused on the relation between self- and peer-informant reports (e.g., see Achenbach, Krukowski, Dumenci, & Ivanova, 2005; Klonsky et al., 2002), another potential source of data on PDs comes from the quantified judgments of clinically experienced observers, such as patients' treating clinicians. The focus of this article is on two PDs for which prior data have shown reasonable convergence across methods and observers—BPD and antisocial personality disorder (APD)—using a third disorder (obsessive–compulsive; OCPD) for discriminant validity. In particular, our goal was to compare dimensional measures of these three disorders using the Shedler–Westen Assessment Procedure (SWAP-200), an instrument designed to quantify personality judgments made by clinically experienced informants with ratings based on the Personality Assessment Inventory (PAI), a self-report measure.

### Quantifying Clinical Observations of Personality

The SWAP-200 was designed for statistical prediction using clinically experienced informants (for further information on the crossing of clinical informants and statistical prediction, see Westen & Weinberger, 2004). Clinically experienced observers sort the 200 items of the Q-sort on the basis of either all available clinical data (e.g., in the course of an ongoing treatment) or on data ascertained using a systematic clinical research interview—the Clinical Diagnostic Interview (CDI; Westen & Muderrisoglu, 2003; Westen & Muderrisoglu, 2006). The CDI differs from

---

Rebekah Bradley, Department of Psychiatry and Behavioral Sciences, Emory University; Mark Hilsenroth, Derner Institute of Advanced Psychological Studies, Adelphi University; Clifford Guarnaccia, Department of Psychology, Emory University; Drew Westen, Department of Psychology and Department of Psychiatry and Behavioral Sciences, Emory University.

Preparation of this article was supported in part by National Institute of Mental Health Grants MH62377 and MH62378 awarded to Drew Westen.

Correspondence concerning this article should be addressed to Drew Westen or Rebekah Bradley, Departments of Psychology and Psychiatry, Emory University, 532 North Kilgo Circle, Atlanta, Georgia 30322. E-mail: dwesten@emory.edu or rbradl2@emory.edu

structured PD interviews in that it does not primarily ask patients to describe themselves (although it does not avoid face-valid questions about behaviors, intentions, or phenomenology, e.g., whether the patient has self-mutilated or thought about suicide). Instead, it asks patients to provide detailed narratives about their symptoms, their education and work history, and their relationship history, focusing on specific examples of emotionally salient experiences. From these data (or from all available clinical data, if the clinician is describing a patient in ongoing treatment), the clinician informant makes judgments about the ways the patient characteristically thinks, feels, regulates impulses and emotions, views the self and others, and behaves in significant relationships, reflected in the placement (ranking) of the items.

The SWAP-200 (and the most recent edition of the instrument, the SWAP-II) has been used in a number of studies. Both the adult and adolescent versions of the instrument predict a range of relevant external criteria, from variables that require little judgment, such as history of suicide attempts and history of psychiatric hospitalizations, to measures of adaptive functioning and developmental and family history variables of potential etiologic significance (e.g., history of first-degree relatives with substance-abuse disorders or childhood history of sexual abuse; Bradley, Jenei, & Westen, 2005; Russ, Heim, & Westen, 2003; Shedler & Westen, 2004; Westen & Shedler, 1999a; Westen, Shedler, Durrett, Glass, & Martens, 2003). The adolescent version of the instrument has also been shown to predict patterns of association with the Child Behavior Checklist (Dutra, Campbell, & Westen, 2004) and measures of attachment status (Nakash-Eisikovits, Dutra, & Westen, 2002). The major limitation of several studies to date has been that the same informant (the treating clinician) has provided both the SWAP data and data on criterion variables.

Thus far, three studies have correlated SWAP data with data from an independent informant. In the first study ( $N = 24$ ), two clinicians independently described the patient using the SWAP-200 from the same CDI interview data (i.e., after conducting or watching it on videotape). This procedure yielded a median interrater reliability coefficient of .80 for SWAP PD scale scores, and these data (aggregated to maximize reliability) correlated just as strongly with PD scale scores based on the treating clinician's independent SWAP-200 assessment of the same patient, who was blind to all interview data (Westen & Muderrisoglu, 2003). Using the same sample, we recently found similarly high correlations (and strong evidence of discriminant validity) for factor-analytically derived *trait* scales using the SWAP-200 (Westen & Muderrisoglu, 2006).

In a second study, conducted by an independent research team (Marin-Avellan, McGauley, Campbell, & Fonagy, 2005), the investigators applied the SWAP-200 to audiotaped Adult Attachment Interviews (Main & Goldwyn, 1985) and chart records on a sample of inpatients at a maximum-security forensic hospital (a method similar to methods for coding psychopathy; Hare et al., 1990). Interrater reliability ( $N = 30$ ) for SWAP-200 PD scale scores was high, with a median or  $r = .91$ . SWAP-200 PD scores correlated more strongly than did SCID-II diagnoses with a range of interpersonal variables rated by nurses on the ward using a 49-item interpersonal circumplex rating scale. SWAP diagnoses, unlike SCID-II diagnoses, were also predictive of patients' index offense (e.g., whether it was violent).

In a third study (Davidson, Obonsawin, Seils, & Patience, 2003), investigators devised a self-report version of the SWAP and examined the relation between self- and clinician-reported SWAP data on a small sample of outpatients ( $N = 23$ ). A substantial divergence between the two was found. However, the investigators did not calculate scale scores using the SWAP manual or any other systematic way of aggregating and correlating self- with clinician-report data. Reanalysis of the data by our research group (unpublished data) found correlations from near zero (OCPD and narcissistic PD) to as high as  $r = .60$  (depressive PD) between clinician reports and self-reports when both were scored using procedures described in the SWAP manual.

### The Present Study

Clearly, more research is necessary on the relation between SWAP scale scores and data provided by informants other than the treating clinician. We chose three PDs for this study: APD, BPD, and OCPD. We chose APD and BPD because they are PDs for which self-report data have proven the most valid (i.e., showing the highest convergence across measures and informants). Unlike disorders such as narcissistic PD, the criteria for BPD and APD include many readily observable behaviors, and BPD and APD patients tend to be forthcoming about their symptoms unless they are motivated not to be (e.g., in forensic settings or custody hearings). We chose OCPD as a comparison diagnosis because prior research suggests not only that it differs substantially from these two Cluster B disorders but also that it tends to be associated with relatively less impaired functioning compared with BPD and APD (e.g., Pfohl & Blum, 1995; Skodol et al., 2002). Thus, we expected OCPD scale scores on the SWAP to correlate negatively with the PAI measures of BPD and APD and related dimensions (e.g., substance abuse).

### Method

#### Sample

Participants included 54 patients seeking outpatient treatment in a community-based training clinic, which is known as a sliding-fee community resource and hence draws from a range of referral sources. The patients in this study included consecutive referrals to the psychodynamic psychotherapy treatment team at the clinic. Each patient who entered treatment with this team was presented with the opportunity to participate in the study, and those presented here include the first 54 patients who agreed to participate. Eight patients declined to participate. Patients were accepted into treatment regardless of disorder or comorbidity. Level of psychological distress and disturbance was primarily in the mild to moderate range, as evidenced by *Diagnostic and Statistical Manual of Mental Disorders*, fourth edition (*DSM-IV*; American Psychiatric Association, 1994) diagnostic categories, clinician rating scales, and self-report measures (Hilsenroth, Ackerman, & Blagys, 2001; Hilsenroth, Ackerman, Blagys, Baity, & Mooney, 2003; Hilsenroth et al., 2000; Hilsenroth, Callahan, & Eudell, 2003).

#### Procedure

Each participant was administered a videotaped systematic clinical interview, derived from the CDI, that lasted approximately 2

hr and an interpretive/feedback interview that lasted approximately 1 hr. The clinical interview focused on presenting problems; psychiatric, medical, family, developmental, social, educational, and work history; an exploration of both past and current relational episodes; and a mental status examination that included an assessment of all *DSM-IV* symptom criteria for schizophrenia, major depression, mania, and dysthymia. The feedback session was organized according to the therapeutic model of assessment (Finn & Tonsager, 1997), aimed at alliance building, exploration of factors maintaining life problems and potential solutions, and examination of therapist-patient interaction. The clinician who conducted the assessment also served as the patient's psychotherapist following the evaluation and provided a SWAP-200 description of the patient after the second session following the feedback session (i.e., the fifth clinical contact hour).<sup>1</sup> A second clinically trained observer independently described the patient using the SWAP-200 based on videotapes of the same five sessions.

Eighteen advanced students (6 men and 12 women) enrolled in an American Psychological Association (APA)-approved clinical psychology PhD program conducted the assessment and psychotherapy sessions and served as coders for the present study. All clinicians had completed graduate coursework in descriptive psychopathology and were supervised by a licensed clinical psychologist with several years of applied experience. Prior to assessing or treating patients for this study, clinicians participated in both individual and group training on scoring for each of several variables of interest, including the SWAP-200. Clinicians also received a minimum of 3.5 hr of supervision per week (1.5 hr individually and 2 hr in a group treatment team meeting) on the therapeutic assessment model/process, scoring/interpretation of assessment measures, clinical interventions, and presentation/organization of collaborative feedback.

### Measures

**SWAP-200.** The SWAP-200 (Westen & Shedler, 1999a, 1999b) is a Q-sort instrument designed to assess personality pathology, including the *DSM-IV* PDs. A Q-sort is a set of statements printed on separate index cards, in this case, statements about personality and personality dysfunction. In the present implementation of the Q-sort method (the SWAP-200), an experienced clinician sorts (rank orders) the statements into eight categories (piles), from those that are not descriptive of the patient (assigned a value of 0) to those that are highly descriptive (assigned a value of 7), with intermediate placement of items that apply in varying degrees. Clinician judges sort the items using a fixed distribution (see Block, 1978). The SWAP-200 thus provides a numeric score ranging from 0 to 7 for each of 200 personality-descriptive items.

SWAP-200 PD scale scores are obtained by correlating patients' 200-item profiles with aggregate profiles of criterion groups from a normative sample and transforming to *T* scores (for more information on the derivation of these scores, see Westen & Shedler, 1999a, 1999b). Thus, a patient's dimensional score for each PD (which we refer to hereafter as a PD score) reflects the degree of match between the evaluating clinician's description of the patient and an empirical prototype. The use of Q correlations of this sort is not novel and has a number of useful psychometric properties (Block, 1971). Like the approach used in the MMPI-2, SWAP-200

PD scores reflect an empirical criterion-keying approach, which diagnoses pathology on the basis of a measure of match between the patient's profile and a normative sample of patients with a given diagnosis (used as a criterion group).

**PAI.** The PAI (Morey, 1991) is a 344-item multiscale inventory assessing clinical syndromes, treatment-related variables, interpersonal style, and response bias. Of particular relevance to the present investigation are the Borderline Features Scale and subscales (Affective Instability, Identity Problems, Negative Relationships, and Self-Harm), the Antisocial Features Scale and subscales (Antisocial Behaviors, Egocentricity, and Stimulus Seeking), the Aggression Scale and subscales (Aggressive Attitudes, Verbal Aggression, and Physical Aggression), the Alcohol and Drug Abuse Scales, and the Suicidal Ideation Scale. Numerous studies support the reliability and validity of the instrument (e.g., Edens, Hart, Johnson, Johnson, & Olver, 2000; Morey, 1991).

## Results

### Demographics

The sample included 23 men and 31 women. Mean age was 30 years ( $SD = 11.35$ ); 28 patients were single, 13 were married, and 13 were divorced. *DSM* diagnoses were assigned by applying *DSM-IV* (American Psychiatric Association, 1994) criteria to data gathered in the systematic clinical interview (described above). *DSM-IV* Axis I diagnoses included mood disorder ( $n = 36$ ), anxiety disorder ( $n = 3$ ), substance-related disorder ( $n = 1$ ), adjustment disorder ( $n = 7$ ), and V-code relational problems ( $n = 6$ ). Axis II diagnoses included borderline ( $n = 1$ ), dependent ( $n = 1$ ), paranoid ( $n = 1$ ), schizotypal ( $n = 1$ ), avoidant ( $n = 1$ ), and PD not otherwise specified (NOS;  $n = 20$ ).

### Relationship Between the SWAP-200 APD, BPD, and OCPD Scores and PAI Scale Scores

Table 1 presents correlations for the SWAP BPD, APD, and OCPD scales with the relevant PAI scales. As would be predicted from prior literature on cross-informant correlations in which one of the informants is the target person (in this case, a patient), the correlation between SWAP scales and PAI scales ranged from small to moderate. The SWAP BPD scale correlated significantly with the overall PAI BPD scale as well as the affect regulation, identity disturbance, and negative relationship BPD subscales. The relationship between the SWAP BPD scale and the PAI Self-Harm scale was not significant. The SWAP APD scale correlated significantly with two of the four antisocial PAI scales (overall antisocial and antisocial behavior) but not with the egocentricity scale or stimulation-seeking scale of the PAI. The pattern of findings also showed clear discrimination between BPD and APD, which are near-neighbor disorders, suggesting discriminant validity.

<sup>1</sup> In five cases, the clinicians completed the SWAP-200 after only the two-session intake interview because the referral was for assessment rather than psychotherapy. In another case, the clinicians completed the SWAP-200 after the first psychotherapy session following the feedback session because of premature termination. These irregularities, however, should diminish the reliability and validity of SWAP diagnoses and hence lead to more conservative findings rather than less conservative findings.

Table 1  
*Predicted and Obtained Correlations Between SWAP BPD, APD, and OCPD Scores and PAI Scales (n = 47)*

PAI scale	SWAP BPD score	SWAP APD score	SWAP OCPD score
<b>BPD</b>			
Borderline Features	.31*	.15	-.35*
Affective Instability	.40**	.18	-.29*
Identity Problems	.33*	.03	-.23
Negative Relationships	.33*	.23	-.36*
Self-Harm	-.07	.04	.17
<b>APD</b>			
Antisocial Features	.09	.35*	-.39**
Behavior	.03	.44**	-.28*
Egocentricity	.07	.12	.08
Stimulation Seeking	.05	.21	-.54**
<b>Aggression</b>			
Aggressive Attitudes	.38**	.45**	-.28*
Verbal Aggression	.20	.42**	-.21
Physical Aggression	.31*	.37**	-.35*
Drug Problems	.28	.46**	-.49**
Alcohol Problems	.23	.14	-.03
Suicidal Ideation	.21	-.12	-.14

Note. SWAP = Shedler-Westen Assessment Procedure; BPD = borderline personality disorder; APD = antisocial personality disorder; OCPD = obsessive-compulsive personality disorder; PAI = Personality Assessment Inventory.

\*  $p < .05$ . \*\*  $p < .01$ .

As predicted, both the SWAP BPD and SWAP APD scales correlated with the PAI aggression scales, which was expected because of the link between each of the disorders and aggression (e.g., impulsive aggression in BPD, impulsive and instrumental aggression in APD). With this sample size, the SWAP BPD scale did not significantly correlate with the PAI suicidal ideation or alcohol scales, although both variables showed trends. As expected, the SWAP OCPD correlated significantly and negatively with most of the PAI BPD and APD scales and subscales as well as with scales measuring aggressive attitudes, physical aggression, and drug problems.

## Discussion

### Primary Findings

The results of this study find small to moderate correlations between personality as assessed by clinician report using the SWAP-200 and personality assessed by self-report using the PAI. These data are consistent with prior literature on self-informant cross-correlations. In general, the SWAP BPD scale correlated with the PAI BPD scales, and the SWAP APD scale correlated with the PAI antisocial scales. As predicted, both the SWAP BPD and SWAP APD scales tended to correlate with PAI scales assessing aggression. Cross-correlations tended to be substantially higher for the same disorder (e.g., SWAP BPD and PAI BPD) than the other disorder, even though the two were from the same Axis II cluster (Cluster B) and have demonstrated high comorbidity in prior research. Further demonstrating discriminant validity were the strong negative correlations between SWAP OCPD scale

scores and the PAI scores related to BDP, APD, aggression, and substance use and abuse.

A notable exception to the expected correlations was the lack of significant association between the SWAP BPD scale and the PAI Self-Harm and Suicidal Ideation scales. This likely reflects the level of functioning of patients in this sample (Global Assessment of Functioning,  $M = 60.98$ ,  $SD = 6.28$ ), who tended, even when they had BPD features, not to make suicidal or parasuicidal gestures. It is of some note that the correlations between clinician report and self-report are similar to those identified by prior studies examining the relationship between lay-observer report and self-report.

### Limitations and Directions for Future Research

The present study has the following limitations. First, clinicians and independent assessors were inexperienced relative to the intended users of the SWAP-200 (experienced clinicians). Second, the sample was relatively high functioning, with a somewhat restricted range of personality pathology. Roughly half met criteria for a *DSM-IV* PD diagnosis, but few were actively self-harming, getting arrested for crimes, and so forth. Third, the data do not address categorical (present/absent) PD diagnoses.

To our knowledge, however, this is the first study to compare SWAP-200 diagnoses with self-report PD instruments (other than the interviewer-rated Structured Clinical Interview of Personality Disorders [SCID-II], on which Marin-Avelon and colleagues found substantial correlations between SWAP-200 and SCID-II dimensional diagnoses). The data provide evidence for convergent and discriminant validity in the range of other cross-informant studies. Whether one method is more or less valid, however, cannot be ascertained from this or any prior study. Clearly, the next step in this research is to collect data on a larger, broader sample of patients using not only the SWAP as applied to CDI data but also other PD instruments widely in use and to assess their relative ability to predict a range of criterion variables indicative of validity of a diagnostic instrument.

## References

- Achenbach, T. M., Krukowski, R. A., Dumenci, L., & Ivanova, M. Y. (2005). Assessment of adult psychopathology: Meta-analyses and implications of cross-informant correlations. *Psychological Bulletin*, *131*, 361–382.
- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- Block, J. (1971). *Lives through time*. Berkeley, CA: Bancroft Books.
- Block, J. (1978). *The Q-sort method in personality assessment and psychiatric research*. Palo Alto, CA: Consulting Psychologists Press.
- Bradley, R., Jenei, J., & Westen, D. (2005). Etiology of borderline personality disorder: Disentangling the contributions of intercorrelated antecedents. *Journal of Nervous and Mental Disease*, *193*, 24–31.
- Clifton, A., Turkheimer, E., & Oltmanns, T. F. (2005). Self- and peer perspectives on pathological personality traits and interpersonal problems. *Psychological Assessment*, *17*, 123–131.
- Davidson, K. M., Obonsawin, M. C., Seils, M., & Patience, L. (2003). Patient and clinician agreement on personality using the SWAP-200. *Journal of Personality Disorders*, *17*, 208–218.
- Dutra, L., Campbell, L., & Westen, D. (2004). Quantifying clinical judgment in the assessment of adolescent psychopathology: Reliability,



- validity, and factor structure of the Child Behavior Checklist for clinician report. *Journal of Clinical Psychology*, 60, 65–85.
- Edens, J. F., Hart, S. D., Johnson, D. W., Johnson, J. K., & Olver, M. E. (2000). Use of the Personality Assessment Inventory to assess psychopathy in offender populations. *Psychological Assessment*, 12, 132–139.
- Finn, S. E., & Tonsager, M. E. (1997). Information-gathering and therapeutic models of assessment: Complementary paradigms. *Psychological Assessment*, 9, 374–385.
- Hare, R. D., Harpur, T. J., Hakstian, A. R., Forth, A. E., Hart, S. D., & Newman, J. P. (1990). The Revised Psychopathy Checklist: Reliability and factor structure. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, 2, 338–341.
- Hilsenroth, M. J., Ackerman, S. J., & Blagys, M. D. (2001). Evaluating the phase model of change during short-term psychodynamic psychotherapy. *Psychotherapy Research*, 11, 29–47.
- Hilsenroth, M. J., Ackerman, S. J., Blagys, M. D., Baity, M. R., & Mooney, M. (2003). Short-term psychodynamic psychotherapy for depression: An evaluation of statistical, clinically significant, and technique-specific change. *Journal of Nervous and Mental Disease*, 191, 349–357.
- Hilsenroth, M. J., Ackerman, S. J., Blagys, M. D., Baumann, B. D., Baity, M. R., Smith, S. R., et al. (2000). Reliability and validity of *DSM-IV* Axis V. *American Journal of Psychiatry*, 157, 1858–1863.
- Hilsenroth, M. J., Callahan, K. L., & Eudell, E. M. (2003). Further reliability, convergent and discriminant validity of overall defensive functioning. *Journal of Nervous and Mental Disease*, 191, 730–737.
- Klein, D. N. (2003). Patients' versus informants' reports of personality disorders in predicting 7 1/2-year outcome in outpatients with depressive disorders. *Psychological Assessment*, 15, 216–222.
- 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 and Practice*, 9, 300–311.
- Main, M., & Goldwyn, R. (1985). *Adult attachment classification system*. Unpublished manuscript, University of California, Berkeley.
- 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.
- Morey, L. C. (1991). *The Personality Assessment Inventory: Professional manual*. Odessa, FL: Psychological Assessment Resources.
- Nakash-Eisikovits, O., Dutra, L., & Westen, D. (2002). Relationship between attachment patterns and personality pathology in adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*, 41, 1111–1123.
- Oltmanns, T. F., Melley, A. H., & Turkheimer, E. (2002). Impaired social functioning and symptoms of personality disorders assessed by peer and self-report in a nonclinical population. *Journal of Personality Disorders*, 16, 437–452.
- Pfohl, B., & Blum, N. (1995). Obsessive-compulsive personality disorder. In W. J. Livesley (Ed.), *The DSM-IV personality disorders* (pp. 261–276). New York: Guilford Press.
- Russ, E., Heim, A., & Westen, D. (2003). Parental bonding and personality pathology assessed by clinician report. *Journal of Personality Disorders*, 17, 522–536.
- Shedler, J., & Westen, D. (2004). Dimensions of personality pathology: An alternative to the five-factor model. *American Journal of Psychiatry*, 161, 1743–1754.
- Skodol, A. E., Gunderson, J. G., McGlashan, T. H., Dyck, I. R., Stout, R. L., Bender, D. S., et al. (2002). Functional impairment in patients with schizotypal, borderline, avoidant, or obsessive-compulsive personality disorder. *American Journal of Psychiatry*, 159, 276–283.
- Thomas, C., Turkheimer, E., & Oltmanns, T. F. (2003). Factorial structure of pathological personality as evaluated by peers. *Journal of Abnormal Psychology*, 112, 81–91.
- 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. (2006). Clinical assessment of pathological personality traits. *American Journal of Psychiatry*, 163, 1285–1297.
- 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., 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., & Weinberger, J. (2004). When clinical description becomes statistical prediction. *American Psychologist*, 59, 595–613.

Received October 21, 2004

Revision received November 30, 2006

Accepted January 9, 2006 ■