Assessing Adolescent Personality Disorders With the Shedler–Westen Assessment Procedure for Adolescents

Jared A. DeFife, Department of Psychiatry and Behavioral Sciences, Emory University School of Medicine

Johanna C. Malone, Department of Psychiatry, Massachusetts General Hospital Harvard Medical School

John DiLallo, New York University School of Medicine, New York City Administration for Children's Services

Drew Westen, Departments of Psychology and Psychiatry, Emory University

This two-part study describes the development and validation of a method for quantifying adolescent personality pathology using the latest edition of the Shedler–Westen Assessment Procedure for Adolescents (SWAP-II-A), an instrument designed to be used by clinically experienced observers. In Study 1, experienced psychologists and psychiatrists described a normative clinical sample of 950 North American patients. Study 2 applied the SWAP-II-A in a day treatment setting. Results indicated that SWAP-II-A personality disorder (PD) scales evidenced high internal consistency, construct validity with *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.) symptoms and diagnoses, and concurrent validity with Child Behavior Checklist (CBCL) ratings. Independent observers saw patients similarly, and PD assessments were significantly associated with CBCL scale scores and ward behavior.

***Key words:* adolescent personality disorders, diagnosis and classification, dimensional diagnosis, personality pathology, Shedler–Westen Assessment Procedure— Adolescent version. [*Clin Psychol Sci Prac* 20: 393–407, 2013]**

Address correspondence to Jared A. DeFife, Department of Psychiatry and Behavioral Sciences, Emory University School of Medicine, 1365 Clifton Rd., NE, Building B, Suite B6117, Atlanta, GA 30322. E-mail: jdefife@emory.edu.

The *Diagnostic and Statistical Manual of Mental Disorders* (*DSM*; American Psychiatric Association, 2000, 2013) offers a cautionary statement regarding the application of personality diagnosis in adolescents. This caution is based largely on concerns about stigmatization (Becker et al., 1999; Bernstein et al., 1993; Hinshaw & Cicchetti, 2000; Mattanah, Becker, Levy, Edell, & McGlashan, 1995) and questions about the stability of personality features through child and adolescent development (Clark, 2009). However, a substantial body of research produced since the initial publication of the *DSM-IV* shows that many aspects of adolescent personality functioning predict adult personality functioning, that personality pathology in the teenage years is clinically distinctive from normative development, and that personality disorders (PDs) are effectively diagnosable in adolescence (Bornovalova, Hicks, Iacono, & McGue, 2009; Donnellan, Conger, & Burzette, 2007; Durrett & Westen, 2005; Levy et al., 1999; McGue, Bacon, & Lykken, 1993; Patricia, 2008; Roberts, Caspi, & Moffitt, 2001; Roberts & DelVecchio, 2000; Skodol, Johnson, Cohen, Sneed, & Crawford, 2007; Westen, Betan, & DeFife, 2011; Westen, Shedler, Durrett, Glass, & Martens, 2003). While many studies demonstrate that adolescent PD diagnosis is related to concurrent impairment on a range of clinical and global

assessment of functioning (GAF) scales including depression, identity disturbance, self/other representations, and interpersonal difficulties (Daley, Rizzo, & Gunderson, 2006; DeFife, Goldberg, & Westen, in press; Levy et al., 1999; Westen et al., 2011), others indicate that adolescent PDs are longitudinally predictive of later functioning and increased vulnerability to mood, anxiety, and substance use disorders and other Axis I conditions (Crawford & Cohen, 2008; Skodol et al., 2007), and can predict Axis I disorders not present in adolescence.

Ongoing questions remain, however, about how best to classify and assess personality pathology in both adults and adolescents (Krueger et al., 2011; Westen, Shedler, Bradley, & DeFife, 2012). Although researchers from multiple perspectives and research traditions disagree on many key questions about personality diagnosis, many agree that dimensional personality diagnosis has substantial advantages over categorical classification in both adults and adolescents (Krueger, 2005; Shedler et al., 2010; Westen, Gabbard, & Blagov, 2006; Westen et al., 2012; Widiger & Samuel, 2005). Dimensional personality diagnosis, however, can take many forms, including dimensionalizing current *DSM* diagnoses (e.g., creating a quantitative score by summing diagnostic criteria rather than imposing an arbitrary categorical cutpoint; Oldham & Skodol, 2000; Zimmerman, Chelminski, Young, Dalrymple, & Martinez, 2011); dimensional trait models derived via factor analysis (e.g., Krueger & Markon, 2011; Krueger, Watson, & Barlow, 2005; Westen, Waller Shedler, & Blagov, in press; Widiger, 2005); diagnosis by prototype matching, in which diagnosticians gauge the similarity or "fit" between a patient's clinical presentation and empirically derived diagnostic prototypes that describe personality syndromes in their "ideal" or prototypical form (e.g., Rottman, Ahn, Sanislow, & Kim, 2009; Spitzer, First, Shedler, Westen, & Skodol, 2008; Westen, 2012; Westen & Shedler, 2000; Westen, Shedler, & Bradley, 2006); or development of empirically derived scales and syndromes using items from existing psychometrically validated assessment measures (e.g., Jones, 2005; Lynam & Widiger, 2001; Miller, Bagby, Pilkonis, Reynolds, & Lynam, 2005; Millon & Davis, 1997; Morey, Waugh, & Blashfield, 1985; Westen & Shedler, 1999a; Westen et al., 2012).

Following this latter approach, the goal of this article was to examine the psychometric development and application of an existing measure of adolescent personality, in this case the Shedler–Westen Assessment Procedure—Adolescent version (SWAP-II-A), for quantifying adolescent personality pathology. The Shedler–Westen Assessment Procedure (SWAP) is a method designed to provide a comprehensive assessment of personality and personality pathology in adolescents and adults (Shedler & Westen, 2004a, 2004b, 2007; Westen & Shedler, 1999a, 1999b, 2007; Westen et al., 2003, 2012). Unlike most personality assessment instruments, which rely primarily or exclusively on patient self-reports (whether by questionnaire or structured interview), the SWAP (which has adult and adolescent versions) is a 200-item Q-sort procedure designed for use by clinically experienced observers in the context of either a thorough examination of a patient using a systemic clinical research interview (DeFife & Westen, 2012; Westen & Muderrisoglu, 2003, 2006) or in a professional assessment or ongoing therapeutic engagement (e.g., longitudinal knowledge of the patient over the course of treatment). To describe a patient using the SWAP, the clinician sorts (rank–orders) the 200 personality statements into eight categories based on their applicability to the patient, from those that are not descriptive (assigned a value of 0) to those that are highly descriptive (assigned a 7). Statements that apply to a greater or lesser degree are placed in intermediate categories.

Shedler–Westen Assessment Procedure for Adolescents data can be analyzed via conventional factor analysis (a variable-centered approach) to identify underlying personality factors or trait dimensions (Westen, Dutra, & Shedler, 2005; Westen et al., in press). They can also be analyzed by factoring patients rather than items via Q-factor analysis (a person-centered approach) to identify groupings of patients who are psychologically similar to one another and distinct from patients in other groupings (i.e., empirically derived personality syndromes). Thus, the SWAP instruments can be used to derive both trait dimensions and naturally occurring diagnostic groupings in the clinical population. Because the item sets for both adult and adolescent versions of the SWAP cover the domains included in the *DSM*, the instruments can additionally be used to derive scales for dimensional assessment of *DSM-5* PD diagnoses, an

approach we investigate below using two independent adolescent clinical samples.

In Study 1, we empirically derived and validated adolescent PD diagnostic criteria and scales using data from a large national sample of practicing clinicians. Study 2 applied these scales to a day treatment sample, in which two independent doctoral-level observers who knew the patients in different contexts assessed them using the SWAP-II-A and a second pair of independent clinical observers independently rated their behavior on the ward using a well-validated instrument, and we used the scales developed in Study 1 to assess cross-observer consistency, reliability, and validity.

STUDY 1

Method

Procedure. As part of a larger study on adolescent personality pathology (Westen, Nakash, Thomas, & Bradley, 2006; Westen & Shedler, 2007), we contacted a random national sample of psychiatrists and psychologists with at least five years of experience postresidency (MDs) or postlicensure (PhDs) from the membership registers of the American Psychiatric and American Psychological Associations, including clinicians targeted in prior solicitations to create a practice research network. We selected clinicians whose membership records indicated an interest in children or adolescents, and supplemented this where necessary with a general sample, given that many clinicians treat adolescents as well as adults. Over one-third of clinicians agreed to participate by the time we had recruited the complete sample, resulting in data on a sample of 950 patients described by their treating clinicians on a range of measures, with psychologists represented at roughly twice the rate as psychiatrists. Participating clinicians received a consulting fee of $200 to complete a battery of measures. Clinicians received a packet containing a cover letter, a consent form, a postage-paid return envelope, and the study measures. Each clinician contributed data on only one patient, to minimize rater-dependent variance.

To obtain a broad range of personality pathology, from relatively minimal to substantial, we asked clinicians to describe "an adolescent patient you are currently treating or evaluating who has enduring patterns of thoughts, feeling, motivation, or behavior—that is, personality problems—that cause distress or dysfunction,"

and emphasized that patients need not have a *DSM* PD diagnosis. We also instructed clinicians to disregard the caveats in the *DSM* regarding the application of Axis II diagnoses to adolescents and simply to select a patient with any degree or form of personality pathology.

We obtained a stratified random sample, stratifying on age (13–18) and sex. The only exclusion criteria were chronic psychosis and mental retardation. In addition, we asked clinicians to select a patient whose personality they felt they knew, using as a guideline ≥6 clinical contact hours but ≤2 years (to minimize confounds imposed by personality change with treatment). To minimize selection biases, we directed clinicians to consult their calendars to select the last patient they saw during the previous week who met study criteria, regardless of setting (e.g., private practice, residential facility). For the purposes of this study, we randomly split the sample to create a derivation sample and a validation sample.

Measures. The core battery of measures required approximately two hours to complete. We describe here only the measures used in this report.

Clinical Data Form for Adolescents (CDF-A). The CDF-A (the adolescent version of the Clinical Data Form; see Westen et al., 2003) is a form designed for use by experienced clinicians developed over several years that assesses a range of variables related to demographics, diagnosis, adaptive functioning, developmental and family history, and other variables of potential etiological significance (Westen & Shedler, 1999a). Research has demonstrated that clinician ratings of adaptive functioning variables show high inter-rater reliability and concurrent validity (e.g., correlations with the same data obtained by independent interview $r > 0.60$; Hilsenroth et al., 2000; Westen, Muderrisoglu, Fowler, Shedler, & Koren, 1997). In prior studies with both adolescent and adult samples, clinicians' judgments on current adaptive functioning and family and developmental history variables have predicted theoretically relevant criterion variables and have reflected reasonable (and conservative) decision rules (Dutra, Campbell, & Westen, 2004; Nakash-Eisikovits, Dutra, & Westen, 2002) as well as demonstrating extremely high correlations (most ranging from $r = 0.50$ to 0.70) between clinician-report and self-report data (DeFife, Drill, Nakash, & Westen, 2010).

Shedler–Westen Assessment Procedure for Adolescents, Version II (SWAP-II-A). The SWAP-II-A, the most recent version of the SWAP for adolescents (Westen et al., 2003, 2005), has been described above. Both the adult and adolescent versions of the SWAP show considerable evidence of reliability and validity (Westen & Shedler, 2007), predicting a range of measures of adaptive functioning (e.g., history of hospitalizations, school performance, violence), psychopathology (e.g., the Child Behavior Checklist [CBCL]), etiological variables (e.g., childhood history of physical and sexual abuse, family history of internalizing and externalizing disorders), and personality as assessed by independent interviewers blind to clinician data (Blagov & Westen, 2008; DiLallo, Jones, & Westen, 2009; Hinrichs, DeFife, & Westen, 2011; Nakash-Eisikovits et al., 2002; Powers & Westen, 2011; Westen & Shedler, 1999a, 1999b; Westen et al., 2003). In studies with adult samples, the SWAP demonstrates excellent cross-observer validity, with correlations between independent observations from research interviewers and treating clinicians in the 0.50–0.90 range across scales, samples, and studies (Westen et al., 2012). Empirically, clinicians' theoretical orientation and professional degree (psychology or psychiatry) have little impact on the way they use the instrument (Shedler & Westen, 2004a, 2004b).

DSM PD Criterion Checklist. Clinicians completed a randomly ordered checklist of diagnostic criteria for all *DSM* PDs, indicating which criteria the patient met. Categorical diagnoses were derived by adding the number of criteria present and applying the *DSM* decision rules (regarding the number of symptoms required) to generate diagnoses. This method provides results that mirror those of structured diagnostic interviews in both adolescents and adults (Blais & Norman, 1997; Morey, 1988; Westen et al., 2003).

Child Behavior Checklist—Clinician Version. The CBCL (Achenbach, 1991a) is a questionnaire designed to assess behavioral problems and social competencies of children aged 4–18. Three versions of the questionnaire can be completed by the parent, teacher, or child (Achenbach, 1991b, 1991c). The parent-report version of the CBCL also has demonstrated validity as rated by an adolescent's treating clinician (Dutra et al., 2004). This study utilized the parent-/clinician-report version (Achenbach, 1991a). The CBCL includes 128 items grouped into 11 *problem scales* including two broadband scales of *internalizing* and *externalizing* symptomatology. The CBCL is widely employed in both clinical and research settings and has demonstrated strong reliability and validity.

Results

In the complete sample, the patients described ($N = 950$) were evenly split by gender (50.6% female). Ethnicity represented included 78.6% identified as Caucasian, with black people equaling 7.8% and Hispanic 7.2%. The sample comprised a roughly normal distribution with respect to social class, with a slight skew toward middle- and upper-middle class. GAF scores indicated clinically significant impairment ($M = 56.8$, $SD = 9.8$). The most common syndromes were unipolar mood, generalized anxiety, attention-deficit, adjustment, and substance use disorders. As assessed by applying *DSM* criteria to the PD Criterion Checklist, antisocial (33%), avoidant (28%), and borderline (23%) PDs were the most prevalent diagnoses, although all *DSM* PDs were represented in relatively high numbers (with the exception of schizotypal PD, present in only 50 cases), with high rates of comorbidity similar to those found in studies using structured interviews with both adolescents and adults; 650 patients (68.4%) met criteria for a *DSM* PD diagnosis. Clinician respondents were highly experienced ($M = 18.5$ years of practice experience, $SD = 8.6$) and diverse in theoretical orientation, with 52.1% self-defining as "integrative-eclectic." Table 1 presents the patient and clinician demographics split by randomized sample (derivation and validation samples). The split-sample randomization (SPSS 18.0) was highly successful, as evidenced by the lack of any significant group differences across demographic and clinical variables despite the large sample sizes in each half-sample.

Derivation Sample. To develop scales for each diagnosis, we first correlated each SWAP-II-A item with the number of criteria met for each PD using the *DSM* PD Criterion Checklist. We then listed items in descending order of magnitude of correlations for each disorder. Decisions about item inclusion and exclusion thresholds in final scales were based on psychometric considerations, taking into account item–scale

	Derivation Sample (n = 481)	Validation Sample (n = 469)
Patient demographics		
Age	15.55 (1.61)	15.58 (1.58)
Sex		
Female (%)	51	51
Male (%)	49	49
Primary Axis I diagnosis		
Dysthymia (%)	39.7	42.0
ADD/ADHD (%)	28.1	26.9
MDD (%)	27.2	27.9
GAD/anxiety NOS (%)	26.6	26.9
Adjustment (%)	26.0	21.8
Substance use (%)	16.3	18.6
Axis II diagnosis		
Paranoid (%)	14.5	12.8
Schizoid (%)	13.0	13.7
Schizotypal (%)	5.5	5.1
Antisocial (%)	34.0	32.3
Borderline (%)	23.3	22.9
Histrionic (%)	10.9	10.7
Narcissistic (%)	17.4	17.3
Avoidant (%)	28.1	29.1
Dependent (%)	7.8	7.5
Obsessive–compulsive (%)	4.8	5.6
Global assessment of functioning (GAF)	57.02 (9.39)	56.55 (10.15)
Clinician demographics		
Discipline		
Psychiatry (%)	29.3	27.4
Psychology (%)	70.7	72.6
Sex		
Female (%)	42.7	42.2
Male (%)	57.3	57.8
Years of experience	18.50 (8.77)	18.48 (8.51)
Theoretical orientation		
Integrative/eclectic (%)	52.0	52.2
Psychodynamic (%)	19.5	17.9
Cognitive behavioral (%)	18.9	22.2
Biological (%)	4.0	2.8
Other (%)	5.6	4.9
Treatment characteristics		
Length of time in treatment (in months)	12.38 (10.01)	12.34 (10.22)
Clinical setting		
Private practice (%)	71.1	69.4
Outpatient clinic (%)	18.5	17.8
Inpatient/residential (%)	6.2	7.1
School (%)	2.7	2.6
Forensic (%)	2.1	2.1

Note. Group comparisons indicated no significant differences between samples (*p* > .05).

correlations for each personality syndrome. As a guiding principle, items were retained for a given diagnostic scale if they correlated ≥0.30 with number of checklist criteria met for the diagnosis, if the corrected item–scale correlation was ≥0.30, and if inclusion of the item did not suppress alpha reliability of the scale (with the goal of maintaining alpha reliabilities ≥0.70). Decisions that fell in gray areas (e.g., items that correlated with *DSM* symptom criteria but had low item–scale

correlations) were resolved conceptually, with items included if they were consistent with the broader constructs. For example, the SWAP item "appears to fear being alone" showed a low item–scale correlation with other dependent PD scale items, but raised coefficient alpha of the scale and is central to the construct of dependent PD and hence was included in the final scale. For interested readers, an online supplement (Appendix S1) is available detailing the items and corrected item–scale correlations for all PD scales along with coefficient alpha for each scale.

Validation Sample. In the validation sample, we first examined the internal consistency of the SWAP-II-A PD scales. The validation sample showed high internal consistency of the scales, with alphas ranging from 0.70 (schizoid PD) to 0.90 (antisocial PD). Next, we examined the degree of diagnostic overlap across scales. Table 2 presents the intercorrelation matrix of SWAP-II-A PD scales with the intercorrelation coefficients of *DSM* symptom totals (number of symptoms per disorder) presented in parentheses for comparison. As can be seen from the table, SWAP-II-A PD scales show similar or appreciably lower cross-disorder correlations than their equivalent *DSM* dimensional diagnoses based on number of symptoms met per disorder (median *r* for SWAP-II-A = −0.04; median *r* for *DSM* symptoms = 0.29).

To assess validity, we examined the relationship of SWAP-II-A PD scales with their *DSM* equivalents in the validation sample. Table 3 presents the correlation of each SWAP-II-A scale with the dimensional symptom total for each *DSM* disorder and the area under the receiver operating curve statistic (AUC) for predicting categorical diagnosis of each disorder. The results show that SWAP-II-A scales are not only highly correlated with dimensional *DSM* diagnosis as typically calculated (number of symptoms per disorder), with a median *r* = 0.61, but also are moderately to highly accurate (Fischer, Bachmann, & Jaeschke, 2003) predictors of categorical disorder diagnosis, with a median AUC = 0.85.

We next examined the correlations between the SWAP-II-A PD scales and CBCL scale scores, as presented in Table 4. We made a priori predictions of the highest expected correlations in boldface, as we did

Table 2. Intercorrelations among Shedler–Westen Assessment Procedure for Adolescents personality disorder scales in the validation sample with intercorrelations among *DSM-5* symptom totals in parentheses (n = 469)

	Paranoid	Schizoid	Schizotypal	Antisocial	Borderline	Histrionic	Narcissistic	Avoidant	Dependent	Obsessive–Compulsive
Paranoid	—									
Schizoid	0.16 (0.37)	—								
Schizotypal	0.29 (0.44)	0.57 (0.64)	—							
Antisocial	0.45 (0.30)	−0.05 (0.13)	−0.03 (0.08)	—						
Borderline	0.29 (0.37)	−0.26 (0.17)	−0.04 (0.30)	0.40 (0.43)	—					
Histrionic	0.12 (0.29)	−0.40 (0.07)	−0.21 (0.19)	0.54 (0.49)	0.60 (0.59)	—				
Narcissistic	0.47 (0.45)	−0.09 (0.17)	−0.05 (0.17)	0.61 (0.54)	0.12 (0.41)	0.45 (0.57)	—			
Avoidant	−0.36 (0.28)	0.39 (0.47)	0.17 (0.45)	−0.67 (−0.16)	−0.46 (0.11)	−0.60 (0.03)	−0.64 (−0.06)	—		
Dependent	−0.40 (0.18)	−0.05 (0.21)	−0.03 (0.30)	−0.51 (0.02)	−0.09 (0.33)	−0.24 (0.26)	−0.58 (0.03)	0.59 (0.50)	—	
Obsessive–compulsive	−0.31 (0.35)	0.08 (0.36)	0.03 (0.37)	−0.55 (0.04)	−0.50 (0.18)	−0.39 (0.16)	−0.12 (0.25)	0.34 (0.36)	0.12 (0.24)	—

Note. Shaded areas indicate correlations between disorders within the same cluster, which are generally expected to overlap.

Table 3. Shedler–Westen Assessment Procedure for Adolescents (SWAP-II-A) personality disorder (PD) scales in the validation sample— Correlations with corresponding *DSM-5* symptom count totals and area under the receiver operating characteristic curve (AUC) for predicting categorical *DSM-5* diagnosis (n = 469)

SWAP-II-A PD Scales	DSM-5 Symptom Counts	AUC
Paranoid	0.49***	0.79***
Schizoid	0.63***	0.86***
Schizotypal	0.57***	0.90***
Antisocial	0.78***	0.90***
Borderline	0.68***	0.84***
Histrionic	0.62***	0.84***
Narcissistic	0.69***	0.87***
Avoidant	0.59***	0.80***
Dependent	0.52***	0.84***
Obsessive–compulsive	0.51***	0.86***

Note. ***$p \leq .001$.

in all subsequent validity analyses, to minimize the possibility of studywise type I error. The pattern of correlations was largely as expected across both specific and higher-order scales. For example, the SWAP-II-A Cluster B scales and paranoid PD correlated with delinquent behavior, aggressive behavior, and externalizing problems. Within Cluster A, schizoid and schizotypal PD both significantly correlated with CBCL social problems, with schizoid PD more strongly associated with withdrawn problems, and with schizotypal PD more associated with thought problems. Cluster C disorders correlated significantly with anxious and internalizing problems.

Discussion

This first study provides large sample derivation and validation support for a dimensional scaling approach to personality pathology in adolescents using the SWAP-II-A. The large national clinical sample was drawn from a range of clinical settings, described patients with personality dysfunction ranging from absent to severe, and utilized clinicians practicing from a range of theoretical orientations, thus rendering it highly representative of the population of interest. Our findings suggest that PD scales derived from a split sample using the SWAP-II-A yield comparable data in the other half-sample, with characteristics comparable to existing *DSM-5* symptom criteria. Not only was reliability high across diagnoses, but comorbidity (cross-correlations of disorders) was substantially lower

Table 4. Correlations between Shedler–Westen Assessment Procedure for Adolescents personality disorder scales and Child Behavior Checklist (CBCL) scale scores in the validation sample (n = 65)

CBCL Scale Scores

	Withdrawn	Somatic Complaints	Anxious/ Depressed	Social Problems	Thought Problems	Attention Problems	Delinquent Behavior	Aggressive Behavior	Sex Problems	Total Problems	Internalizing Problems	Externalizing Problems
Paranoid	0.30*	0.15	−0.05	0.21	−0.04	0.30**	0.35**	**0.54*****	0.07	0.47***	0.13	0.50***
Schizoid	**0.48*****	0.10	0.16	**0.51*****	0.10	0.23	0.01	−0.02	−0.12	0.26*	0.30*	−0.01
Schizotypal	0.22	0.20	0.16	**0.52*****	**0.40*****	0.45***	0.12	0.25*	0.11	0.46***	0.24*	0.21
Antisocial	−0.11	−0.15	−0.39***	0.13	−0.17	0.37***	**0.58*****	**0.62*****	0.10	0.33**	−0.29*	**0.66*****
Borderline	−0.10	0.07	−0.02	0.16	−0.05	0.23	0.39***	0.54***	0.11	**0.40*****	−0.03	0.51***
Histrionic	−0.40***	−0.24	−0.23	−0.07	−0.09	0.05	0.28*	0.36**	0.33*	0.07	−0.35**	0.35**
Narcissistic	−0.23	−0.24	−0.46***	0.06	−0.15	0.03	0.21	**0.47*****	0.08	0.07	−0.40***	0.40***
Avoidant	**0.43*****	0.13	**0.41*****	0.02	0.10	−0.17	−0.33**	−0.56***	−0.15	−0.13	**0.42*****	−0.51***
Dependent	0.20	**0.02**	**0.42*****	0.06	0.16	−0.07	−0.10	−0.28*	0.07	0.09	0.31**	−0.24
Obsessive– compulsive	0.11	0.21	**0.36****	−0.04	0.21	−0.22	−0.53***	−0.54***	−0.16	−0.27*	0.30	−0.57***

Note. Strongest predicted correlations noted in boldface. *p ≤ .05, **p ≤ .01, ***p ≤ .001.

than the most common dimensional measure of PD in adolescents, number of criteria met per disorder.

This reduced comorbidity is a substantial advantage of the SWAP diagnostic assessment approach; however, it could easily have come at a cost to validity, because reduced comorbidity could suggest that we were not assessing the same constructs as in *DSM-5*. That turned out not to be the case as not only were SWAP-II-A PD scales highly related to both dimensional and categorical *DSM* PD diagnoses (evidenced by both correlations and AUC in ROC analyses), but they also associated with CBCL scale scores in predicted ways.

As comorbidity problems have plagued *DSM* diagnosis (Oldham et al., 1992; Zimmerman, Rothschild, & Chelminski, 2005), the SWAP-II-A appears to offer an alternative for researchers and practitioners looking to retain the merits of personality pathology description while increasing diagnostic specificity. Furthermore, use of the SWAP-II-A offers additional benefits, including the availability of empirically derived diagnostic classifications, personality trait scales, and indices of adaptive personality strengths (Westen & Shedler, 2007).

The major limitation of Study 1 is that a single informant (the treating clinician) provided all the data for each case. However, several considerations help to mitigate this concern. First, this limitation is the norm in psychiatric research, and in PD research in particular, in which a single informant (usually the patient) provides all or most of the data (either by self-report or by structured interviews that rely primarily on self-report). The reliance on self-report data is particularly pervasive in personality research (Robins, Tracy, & Sherman, 2007; Schwarz, 1999), where publication surveys of major personality research journals indicate that 95–98% of the articles published are based on data obtained from self-report measurements of personality, with over 70% of cases where self-report instruments were the only measure used (Kagan, 2007; Vazire, 2006). Second, the SWAP instruments (for adults and adolescents) have been widely used in prior research and demonstrate high reliability and cross-observer validity (Blagov, Bi, Shedler, & Westen, 2012; Westen & Muderrisoglu, 2003; Westen et al., 2012) across independent observers. Although clinicians in this study had obtained familiarity with the adolescent patients in their care over the course of treatment, prior studies

suggest that SWAP raters can achieve good to excellent inter-rater reliability ($rs = 0.7$–0.9) following a single-session systematic clinical research interview (Westen & Muderrisoglu, 2003) and that such ratings evidence substantial cross-method correlations with independent measurements obtained from treating clinicians (Westen & Muderrisoglu, 2003; Westen et al., 2012). Finally, the use of the SWAP-II-A in this study is highly reflective of diagnostic procedures in actual clinical practice, where a single clinician uses the best available data (including data from collateral sources, such as parents, teachers, and official records obtained from schools, forensic facilities, etc.) and his or her assessment of the ways the patient functions affectively, cognitively, and interpersonally to formulate a case.

Nevertheless, evidence of cross-observer diagnostic validity is important for further validating the SWAP-II-A PD scales. The goal of Study 2 is to address this limitation of Study 1.

STUDY 2

To examine the cross-observer validity of the SWAP-II-A PD scales, we next assessed the extent to which two doctoral-level observers could independently provide similar ratings on SWAP-II-A PD descriptions and whether these SWAP-II-A PD ratings would correlate in expected ways with other well-validated instruments of psychopathology and ward behavior.

Methods

Participants. Participants were 33 patients from an intensive urban day treatment program. The program was designed for adolescents aged 14–20 from the New York City public school system whose psychiatric problems had interfered with their ability to attend school for a significant period of time. All patients in the program met Medicaid criteria for severely emotionally disturbed (SED), indicated by an initial GAF score of 50 or less. Clinicians included licensed permanent staff as well as mental health trainees from several disciplines.

Procedures. Participants and their guardians (where necessary) were approached individually to solicit consent. Because the SWAP-II-A requires clinically experienced observers, only participants who were patients of licensed clinical staff were selected for this study. Participants were recruited who had been known to their clinicians for at least three months. For each adolescent, two clinicians independently described the patient using multiple instruments: (a) the primary psychotherapist, who was either a licensed clinical psychologist (PhD or PsyD) or a licensed clinical social worker (LICSW) with at least two years of postlicensure experience, and (b) a board-certified child and adolescent psychiatrist, who had at least six years of clinical experience postresidency training. Two other licensed clinicians other than the primary therapist or psychiatrist provided ward behavior ratings (Chart of Interpersonal Reactions in Closed Living Environments [CIRCLE], described below) based on their experience of the adolescent from monitoring the program milieu. Beyond the adolescent completing the process of informed consent, the study method did not require the active participation of patients beyond what had already occurred independently during the course of their treatment.

Measures. *SWAP-II-A.* The SWAP-II-A was administered independently by two experienced clinicians per patient.

Child Behavior Checklist—Clinician Version. The CBCL clinician-report version (Achenbach, 1991a) was used as described in Study 1 above and also completed independently by two experienced doctoral-level clinicians per patient.

Clinical Data Form for Adolescents (CDF-A). The CDF-A was administered independently by two experienced clinicians on each patient.

Chart of Interpersonal Reactions in Closed Living Environments. The CIRCLE (Blackburn & Glasgow, 2006; Blackburn & Renwick, 1996) is a 51-item questionnaire originally designed for nurses to describe the interpersonal verbal and nonverbal behavior of patients using the interpersonal circumplex model (IPC; Leary, 1957). The measure yields eight scales, describing observed behavioral patterns: dominant, coercive, hostile, withdrawn, submissive, compliant, nurturant, and gregarious. Two-week retest correlations ranging from 0.83 to 0.92 demonstrate the short-term stability of the CIRCLE (Blackburn & Renwick, 1996). In addition, the measure has shown meaningful associations with personality pathology in adults (Marin-Avellan,

McGauley, Campbell, & Fonagy, 2005) as well as five-factor personality traits (Blackburn & Glasgow, 2006). CIRCLE data were drawn from two different independent clinical observers who were blind to the assessment data for the instruments described above.

Results

We applied the SWAP-II-A PD scales identified in Study 1 to the SWAP-II-A data provided by two independent raters, with cross-observer intraclass correlation coefficients presented in Table 5. The ratings show high cross-observer agreement (median ICC = 0.60) and are particularly robust in light of the fact that they were obtained from *independent clinical data sources* and not from the same videotaped case material (as would be the case with a more circumscribed inter-rater reliability measurement). To increase reliability for subsequent analyses, we created mean scores across the two clinical raters and used the averaged scores (median ICC = 0.75) for the remaining analyses.

Table 6 presents the correlations between the SWAP-II-A PD scales and CBCL scale scores for the Study 2 sample. The pattern of correlations was largely as expected and similar to the results found in the validation sample for Study 1. Given the small-N sample, the large effect sizes and independent replication engender greater confidence in the results.

Table 7 reports the cross-observer correlations between SWAP-II-A PD scales and blinded independent ratings of interpersonal behavior using the CIRCLE. Once again, the PD scales correlated in expected ways even across these independent rating methods and observers. For example, as predicted, SWAP avoidant

Table 5. Cross-observer ICCs for Shedler–Westen Assessment Procedure for Adolescents *DSM-5* personality disorder scores from an adolescent intensive day treatment program (*n* = 32)

	ICC (1, 2)
Paranoid	0.37 (0.54)
Schizoid	0.81 (0.90)
Schizotypal	0.85 (0.92)
Antisocial	0.69 (0.81)
Borderline	0.43 (0.60)
Histrionic	0.55 (0.71)
Narcissistic	0.64 (0.78)
Avoidant	0.64 (0.78)
Dependent	0.44 (0.61)
Obsessive–compulsive	0.50 (0.67)

Table 6. Correlations between Shedler–Westen Assessment Procedure for Adolescents personality disorders and Child Behavior Checklist (CBCL) scale scores in an intensive day treatment program (*n* = 32)

CBCL Scale Scores

	Withdrawn	Somatic Complaints	Anxious/ Depressed	Social Problems	Thought Problems	Attention Problems	Delinquent Behavior	Aggressive Behavior	Sex Problems	Total Problems	Internalizing Problems	Externalizing Problems
Paranoid	−0.25	−0.55***	−0.25	0.11	−0.09	0.06	0.37*	**0.55***	0.01	0.15	−0.38*	0.51**
Schizoid	**0.64***	0.12	0.01	**0.29**	0.48**	0.28	−0.18	−0.25	−0.17	0.11	0.28	−0.23
Schizotypal	0.44*	0.07	0.01	**0.51**	**0.73***	0.52**	−0.20	−0.14	−0.16	0.21	0.20	−0.16
Antisocial	−0.49**	−0.57**	−0.54**	0.19	−0.34	0.14	**0.92***	**0.80***	0.34	0.25	−0.65**	**0.89***
Borderline	−0.27	−0.09	0.07	0.04	−0.11	0.05	0.10	**0.41***	0.24	**0.22**	−0.11	0.30
Histrionic	−0.54**	−0.09	−0.32	0.03	−0.42*	−0.02	**0.48**	**0.56***	0.22	0.11	−0.43*	**0.56***
Narcissistic	−0.38*	−0.27	−0.54***	−0.10	−0.44**	−0.08	**0.69***	**0.57***	0.28	0.08	−0.53***	**0.64***
Avoidant	**0.70***	0.37*	**0.58***	0.02	0.44*	0.02	−0.74***	−0.74***	−0.31	−0.07	**0.69***	−0.76***
Dependent	0.26	**0.59***	**0.51***	0.38*	0.11	0.27	−0.42*	−0.35	−0.04	0.22	**0.54***	−0.39*
Obsessive–compulsive	0.40*	0.26	**0.26**	−0.30	0.25	−0.29	−0.64***	−0.60***	−0.45**	−0.35	0.37*	−0.65***

Note. Strongest predicted correlations noted in boldface. *p ≤ .05, **p ≤ .01, ***p ≤ .001.

Table 7. Study 2 cross-observer correlations between Shedler–Westen Assessment Procedure for Adolescents personality disorders and Chart of Interpersonal Reactions in Closed Living Environments (CIRCLE) scale scores in an intensive day treatment program (*n* = 32)

CIRCLE Scale Scores

	Dominant	Coercive	Hostile	Withdrawn	Submissive	Compliant	Friendly	Sociable
Paranoid	−0.08	0.22	**0.62*****	0.15	0.01	−0.36*	−0.54*****	−0.30
Schizoid	−0.55*****	−0.42*	0.09	**0.49****	0.45**	0.34	−0.35	−0.50**
Schizotypal	−0.35	−0.32	−0.02	**0.35***	0.31	0.14	−0.15	−0.25
Antisocial	0.18	**0.44***	**0.47****	−0.16	−0.21	−0.33	−0.57**	0.01
Borderline	0.33	**0.34**	0.26	−0.32	−0.38*	−0.38*	0.11	0.25
Histrionic	0.55*****	**0.57*****	0.33	−0.57*****	−0.56*****	−0.53**	−0.01	**0.41***
Narcissistic	**0.37***	**0.47****	0.31	−0.25	−0.39*	−0.42*	−0.30	0.10
Avoidant	−0.64*****	−0.72*****	−0.38*	**0.59*****	**0.64*****	**0.61*****	0.13	−0.42*
Dependent	−0.12	−0.18	−0.18	0.18	0.16	**0.11**	**0.27**	0.07
Obsessive–compulsive	−0.21	−0.42*	−0.41*	0.26	0.17	**0.36***	**0.23**	−0.17

Note. Strongest predicted correlations noted in boldface. *$p \leq .05$, **$p \leq .01$, ***$p \leq .001$.

PD correlated with CIRCLE withdrawn, submissive, and compliant subscales, and antisocial PD correlated with CIRCLE coercive and hostile subscales.

Discussion

In contrast to Study 1, which used data from a single informant, Study 2 utilized ratings from two experienced clinicians describing the same adolescent patient in the context of an intensive day treatment program and two additional informants describing the patients' ward behavior. This study addressed the major limitation of the validity data from Study 1. The cross-observer correlations revealed that two clinicians observing the patients at different times and in different ways produce highly similar SWAP-II-A descriptions of the same patient. These cross-observer validity correlations are substantially stronger than those found for most measures of personality pathology, which tend to average around *r* = 0.30 (Klonsky, Oltmanns, & Turkheimer, 2002). By comparison, the magnitude of our cross-observer coefficients would fall into the highest range (upper quartile) of coefficients commonly observed across standard psychological and medical assessment procedures (Hemphill, 2003).

We were also able to demonstrate convergent and discriminant validity using measures of psychopathology (the CBCL) and ward behavior (the CIRCLE). Even with a small sample size, these cross-observer ratings showed the expected pattern of results and suggest that these personality dimensions correspond with what is seen both in one-on-one interactions and in ward behavior observed by milieu staff in a day treatment setting.

GENERAL DISCUSSION

Together, these findings suggest that the SWAP-II-A offers an effective, pragmatic, reliable, and valid way of quantifying clinical judgment about personality pathology. We used a large-scale clinical sample to derive SWAP-II-A dimensional scales designed to reflect existing *DSM* PD categories. Using an independent validation sample, we found that these SWAP-II-A PD scales demonstrated reduced diagnostic redundancy and substantial construct validity with *DSM* symptom counts, diagnostic efficiency vis-à-vis categorical PD diagnosis, and concurrent validity with common adolescent clinical behavior problems as assessed by the CBCL.

To address concerns about the limitation of single-informant data collection, we followed with a cross-observer validation design using an applied clinical sample in Study 2. SWAP-II-A PD scales showed substantive cross-observer consistency. Not only were concurrent validity results replicated with the CBCL, but the results also showed cross-method validity with ward behavior ratings from independent clinical observers. Although cross-method data from larger and more diverse clinical and nonclinical samples are needed, across the two studies our results show highly promising evidence of validity of the SWAP-II-A PD scales for assessment both within and across clinical observers.

Most importantly, our results suggest that the SWAP-II-A offers a tenable alternative or adjunct to

other personality assessment systems, particularly those that over-rely on self-report data. Both meta-analytic investigations (Klonsky et al., 2002) and data from large-N studies (Clifton, Turkheimer, & Oltmanns, 2003) have shown that self-reported traits and symptoms correlate only moderately with the same traits and symptoms assessed by lay informants (in meta-analytic research, $r = 0.36$) and that self-reported PDs correlate weakly with longitudinal evaluation using all available data (Klein, Ouimette, Kelly, Ferro, & Riso, 1994; Pilkonis, Heape, Ruddy, & Serrao, 1991). With adolescents, where some studies and methods do draw on more than one observer, integrating data across observers can be difficult, as patients and their parents, the most common sources of data, often provide highly divergent data. One of the advantages of the SWAP-II-A is that it provides dimensional scores for each PD that reflect the clinician's use of all available data from all available informants. Further, SWAP-based dimensional PD scale scores from treating clinicians correlate highly with SWAP and other data obtained from other sources, such as independent clinical research interviews (Westen & Muderrisoglu, 2003; Westen et al., 2012). This is, however, the first study to demonstrate this for *DSM* PD dimensional diagnoses in adolescents. The SWAP-II-A offers researchers and clinical practitioners a diagnostic and measurement method that capitalizes on both clinical expertise and sophisticated psychometrics that can be used in both research and practice.[1]

NOTE

1. As with most multidimensional psychopathology measures, linear T-scores for disorders with different base rates can yield very different percentile ranks across disorders in normative samples (Hsu, 1984), creating an obstacle for accurate clinical interpretation. In addressing this psychometric concern for the SWAP instrument (see Blagov et al., 2012), we have generated normalized T-scores (Box & Cox, 1964; Colligan, Osborne, & Offord, 1980, 1984; Waller, 2011), which allows for improved diagnostic interpretation across disorders, irrespective of disorder base rates in the clinical population. Syntax is available on request from the first author.

REFERENCES

Achenbach, T. (1991a). *Manual for the Child Behavior Checklist/ 4-18 and 1991 profile*. Burlington, VT: Department of Psychiatry, University of Vermont.

Achenbach, T. (1991b). *Manual for the teacher's report form and 1991 profile*. Burlington, VT: Department of Psychiatry, University of Vermont.

Achenbach, T. (1991c). *Manual for the youth self-report*. Burlington, VT: Department of Psychiatry, University of Vermont.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Becker, D. F., Grilo, C. M., Morey, L. C., Walker, M. L., Edell, W. S., & McGlashan, T. H. (1999). Applicability of personality disorder criteria to hospitalized adolescents: Evaluation of internal consistency and criterion overlap. *Journal of the American Academy of Child and Adolescent Psychiatry*, *38*, 200–205. doi:10.1097/00004583-19990-2000-00020

Bernstein, D. P., Cohen, P., Velez, C. N., Scwab-Stone, M., Seiver, L. J., & Shinsato, L. (1993). Prevalence and stability of DSM-II-R personality disorders in a community based survey of adolescents. *American Journal of Psychiatry*, *150*, 1237–1243.

Blackburn, R., & Glasgow, D. V. (2006). *Manual for the Chart of Interpersonal Reactions in Closed Living Environments (CIRCLE)*. Unpublished manuscript, University of Liverpool.

Blackburn, R., & Renwick, S. J. (1996). Rating scales for measuring the interpersonal circle in forensic psychiatric patients. *Psychological Assessment*, *8*, 76–84. doi:10.1037/1040-3590.8.1.76

Blagov, P. S., Bi, W., Shedler, J., & Westen, D. (2012). The Shedler–Westen Assessment Procedure (SWAP): Evaluating psychometric questions about its reliability, validity, and impact of its fixed score distribution. *Assessment*, *19*, 370–382. doi:10.1177/1073191112436667

Blagov, P., & Westen, D. (2008). Questioning the coherence of histrionic personality disorder: Borderline and hysterical personality subtypes in adults and adolescents. *Journal of Nervous and Mental Disease*, *196*, 785–797. doi:10.1097/NMD.0b013e31818b502d

Blais, M., & Norman, D. (1997). A psychometric evaluation of the DSM-IV personality disorder criteria. *Journal of Personality Disorders*, *11*, 168–176. doi:10.1521/pedi.1997.11.2.168

Bornovalova, M. A., Hicks, B. M., Iacono, W. G., & McGue, M. (2009). Stability, change, and heritability of borderline personality disorder traits from adolescence to

adulthood: A longitudinal twin study. *Development and Psychopathology*, *21*, 1335–1353. doi:10.1017/S0954579409990186

Box, G. E. P., & Cox, D. R. (1964). Analysis of transformations. *Journal of the Royal Statistical Society, Series B*, *26*, 211–252.

Clark, L. A. (2009). Stability and change in personality disorder. *Current Directions in Psychological Science*, *18*(1), 27–31. doi:10.1111/j.1467-8721.2009.01600.x

Clifton, A., Turkheimer, E., & Oltmanns, T. F. (2003). *Self and peer perspectives on pathological personality traits and interpersonal problems*. Unpublished manuscript, University of Virginia.

Colligan, R., Osborne, D., & Offord, K. (1980). Linear transformation and the interpretation of MMPI T-scores. *Journal of Clinical Psychology*, *36*, 162–165. doi:10.1002/1097-4679(198001)36:1%3c162:AID-JCLP2270360117%3e3.0.CO;2-E

Colligan, R., Osborne, D., & Offord, K. (1984). Normalized transformations and the interpretation of MMPI T-scores: A reply to Hsu. *Journal of Consulting and Clinical Psychology*, *52*, 824–826. doi:10.1037/0022-006X.52.5.824

Crawford, T., & Cohen, P. (2008). Comorbid Axis I and Axis II disorders in early adolescence: Outcomes 20 years later. *Archives of General Psychiatry*, *65*, 641–648. doi:10.1001/archpsyc.65.6.641

Daley, S. E., Rizzo, C. J., & Gunderson, B. H. (2006). The longitudinal relation between personality disorder symptoms and depression in adolescence: The mediating role of interpersonal stress. *Journal of Personality Disorders*, *20*, 352–368. doi:10.1521/pedi.2006.20.4.352

DeFife, J. A., Drill, R., Nakash, O., & Westen, D. (2010). Agreement between clinician and patient ratings of adaptive functioning and developmental history. *American Journal of Psychiatry*, *167*, 1472–1478. doi:10.1176/appi.ajp.2010.09101489

DeFife, J. A., Goldberg, M. G., & Westen, D. (in press). Dimensional assessment of self and interpersonal functioning in adolescents: Implications for DSM-5′s general definition of personality disorder. *Journal of Personality Disorders*. doi:10.1521/pedi_2013_27_085

DeFife, J. A., & Westen, D. (2012). Empirically-informed clinical interviewing for personality disorders. In R. Levy, J. Ablon, & H. Kächele (Eds.), *Psychodynamic psychotherapy research: Evidence-based practice and practice-based evidence* (pp. 553–569). New York, NY: Springer. doi:10.1007/978-1-60761-792-1_31

DiLallo, J. D., Jones, M., & Westen, D. (2009). Personality subtypes in disruptive adolescent males. *Journal of Nervous and Mental Disease*, *197*, 15–23. doi:10.1097/NMD.0b013e318192770c

Donnellan, M. B., Conger, R. D., & Burzette, R. G. (2007). Personality development from late adolescence to young adulthood: Differential stability, normative maturity, and evidence for the maturity-stability hypothesis. *Journal of Personality*, *75*, 237–264. doi:10.1111/j.1467-6494.2007.00438.x

Durrett, C., & Westen, D. (2005). The structure of Axis II pathology in adolescents: A cluster- and factor-analytic investigation. *Journal of Personality Disorders*, *19*, 440–461, doi:10.1521/pedi.2005.19.4.440

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. doi:10.1002/jclp.10234

Fischer, J. E., Bachmann, L. M., & Jaeschke, R. (2003). A readers' guide to the interpretation of diagnostic test properties: Clinical example of sepsis. *Intensive Care Medicine*, *29*, 1043–1051. doi:10.1007/s00134-003-1761-8

Hemphill, J. (2003). Interpreting the magnitudes of correlation coefficients. *American Psychologist*, *58*, 78–79. doi:10.1037/0003-066X.58.1.78

Hilsenroth, M. J., Ackerman, S. J., Blagys, M. D., Baumann, B. D., Baity, M. R., Smith, S. R., … Holdwick, D.J. (2000). Reliability and validity of DSM-IV Axis V. *American Journal of Psychiatry*, *157*, 1858–1863. doi:10.1176/appi.ajp.157.11.1858

Hinrichs, J., DeFife, J. A., & Westen, D. (2011). Personality subtypes in adolescent and adult children of alcoholics: A two-part study. *Journal of Nervous and Mental Disease*, *199*, 487–498. doi:10.1097/NMD.0b013e3182214268

Hinshaw, S. P., & Cicchetti, D. (2000). Stigma and mental disorder: Conceptions of illness, public attitudes, personal disclosure, and social policy. *Development and Psychopathology*, *12*, 555–598. doi:10.1017/S0954579400004028

Hsu, L. M. (1984). MMPI T scores: Linear versus normalized. *Journal of Consulting and Clinical Psychology*, *52*, 821–823. doi:10.1037/0022-006x.52.5.821

Jones, A. (2005). An examination of three sets of MMPI-2 personality disorder scales. *Journal of Personality Disorders*, *19*, 370–385. doi:10.1521/pedi.2005.19.4.370

Kagan, J. (2007). A trio of concerns. *Perspectives on Psychological Science*, *2*, 361–376. doi:10.1111/j.1745-6916.2007.00049.x

Klein, D. N., Ouimette, P. C., Kelly, H. S., Ferro, T., & Riso, L. P. (1994). Test-retest reliability of team consensus best-estimate diagnoses of Axis I and II

disorders in a family study. *American Journal of Psychiatry*, *151*, 1043–1047.

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. doi:10.1093/clipsy.9.3.300

Krueger, R. F. (2005). Continuity of Axes I and II: Toward a unified model of personality, personality disorders, and clinical disorders. *Journal of Personality Disorders*, *19*, 233–261. doi:10.1521/pedi.2005.19.3.233

Krueger, R. F., Eaton, N. R., Clark, L. A., Watson, D., Markon, K. E., Derringer, J., … Livesley, W. J. (2011). Deriving an empirical structure of personality pathology for DSM-5. *Journal of Personality Disorders*, *25*, 170–191. doi:10.1521/pedi.2011.25.2.170

Krueger, R. F., & Markon, K. E. (2011). A dimensional-spectrum model of psychopathology: Progress and opportunities. *Archives of General Psychiatry*, *68*, 10–11. doi:10.1001/archgenpsychiatry.2010.188

Krueger, R. F., Watson, D., & Barlow, D. H. (2005). Introduction to the special section: Toward a dimensionally based taxonomy of psychopathology. *Journal of Abnormal Psychology*, *114*, 491–493. doi:10.1037/0021-843X.114.4.491

Leary, T. (1957). *Interpersonal diagnosis of personality: A functional theory and methodology for personality evaluation*. New York, NY: Ronald Press.

Levy, K. N., Becker, D. F., Grilo, C. M., Mattanah, J., Garnet, K. E., Quinlan, D. M., … McGlashan, T.H. (1999). Concurrent and predictive validity of the personality disorder diagnosis in adolescent patients. *American Journal of Psychiatry*, *156*, 1522–1528.

Lynam, D. R., & Widiger, T. A. (2001). Using the five-factor model to represent the DSM–IV personality disorders: An expert consensus approach. *Journal of Abnormal Psychology*, *110*, 401–412. doi:10.1037/0021-843X.110.3.401

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. doi:10.1002/cbm.35

Mattanah, J. J., Becker, D. F., Levy, K. N., Edell, W. S., & McGlashan, T. H. (1995). Diagnostic stability in adolescents followed up 2 years after hospitalization. *American Journal of Psychiatry*, *152*, 889–894.

McGue, M., Bacon, S., & Lykken, D. T. (1993). Personality stability and change in early adulthood: A behavioral genetic analysis. *Developmental Psychology*, *29*, 96–109. doi:10.1037/0012-1649.29.1.96

Miller, J. D., Bagby, R. M., Pilkonis, P. A., Reynolds, S. K., & Lynam, D. R. (2005). A simplified technique for scoring DSM-IV personality disorders with the five-factor model. *Assessment*, *12*, 404–415. doi:10.1177/1073191105280987

Millon, T., & Davis, R. (1997). *Millon Clinical Multiaxial Inventory-III*. Minneapolis, MN: National Computer Systems.

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., Waugh, M. H., & Blashfield, R. K. (1985). MMPI scales for DSM-III personality disorders: Their derivation and correlates. *Journal of Personality Assessment*, *49*, 245–251. doi:10.1207/s15327752jpa4903_5

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. doi:10.1097/00004583-200209000-00012

Oldham, J. M., & Skodol, A. E. (2000). Charting the future of Axis II. *Journal of Personality Disorders*, *14*, 17–29. doi:10.1521/pedi.2000.14.1.17

Oldham, J., Skodol, A., Kellman, H., Hyler, S., Rosnick, L., & Davies, M. (1992). Diagnosis of DSM-III-R personality disorders by two structured interviews: Patterns of comorbidity. *American Journal of Psychiatry*, *149*, 213–220.

Patricia, C. (2008). Child development and personality disorder. *Psychiatric Clinics of North America*, *31*, 477–493. doi:10.1016/j.psc.2008.03.005

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. doi:10.1037/1040-3590.3.1.46

Powers, A. D., & Westen, D. (2011). Personality subtypes in adolescents with panic disorder. *Clinical Child Psychology and Psychiatry*, *16*, 551–565. doi:10.1177/1359104510387572

Roberts, B. W., Caspi, A., & Moffitt, T. E. (2001). The kids are alright: Growth and stability in personality development from adolescence to adulthood. *Journal of Personality and Social Psychology*, *81*, 670–683. doi:10.1037/0022-3514.81.4.670

Roberts, B. W., & DelVecchio, W. F. (2000). The rank-order consistency of personality traits from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin*, *126*, 3–25. doi:10.1037/0033-2909.126.1.3

Robins, R. W., Tracy, J. L., & Sherman, J. W. (2007). What kinds of methods do personality psychologists use? In R. W. Robins, R. C. Fraley, & R. F. Krueger (Eds.),

Handbook of research methods in personality psychology (pp. 673–678). London, UK: Guilford Press.

Rottman, B., Ahn, W., Sanislow, C., & Kim, N. (2009). Can clinicians recognize DSM-IV personality disorders from five-factor model descriptions of patient cases? *American Journal of Psychiatry*, *166*, 427–433. doi:10.1176/appi.ajp.2008.08070972

Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist*, *54*, 93–105. doi:10.1037/0003-066X.54.2.93

Shedler, J., Beck, A., Fonagy, P., Gabbard, G., Gunderson, J., Kernberg, O., … Westen, D. (2010). Personality disorders in DSM-5. *American Journal of Psychiatry*, *167*, 1026–1028. doi:10.1176/appi.ajp.2010.10050746

Shedler, J., & Westen, D. (2004a). Dimensions of personality pathology: An alternative to the five factor model. *American Journal of Psychiatry*, *161*, 1743–1754. doi:10.1176/appi.ajp.161.10.1743

Shedler, J., & Westen, D. (2004b). Refining personality disorder diagnoses: Integrating science and practice. *American Journal of Psychiatry*, *161*, 1350–1365. doi:10.1176/appi.ajp.161.8.1350

Shedler, J., & Westen, D. (2007). The Shedler–Westen Assessment Procedure (SWAP): Making personality diagnosis clinically meaningful. *Journal of Personality Assessment*, *89*(1), 41–55. doi:10.1080/00223890701357092

Skodol, A. W. E., Johnson, J. G., Cohen, P., Sneed, J. R., & Crawford, T. N. (2007). Personality disorder and impaired functioning from adolescence to adulthood. *British Journal of Psychiatry*, *190*, 415–420. doi:10.1192/bjp.bp.105.019364

Spitzer, R., First, M., Shedler, J., Westen, D., & Skodol, A. (2008). Clinical utility of five dimensional systems for personality diagnosis: A "consumer preference" study. *Journal of Nervous and Mental Disease*, *196*, 356–374. doi:10.1097/NMD.0b013e3181710950

Vazire, S. (2006). Informant reports: A cheap, fast, and easy method for personality assessment. *Journal of Research in Personality*, *40*, 472–481. doi:10.1016/j.jrp.2005.03.003

Waller, N. (2011). *Choosing a scale for SWAP-II factor scores: The relative merits of linear, normalized, and uniform T-scores*. Unpublished manuscript, University of Minnesota.

Westen, D. (2012). Prototype diagnosis of psychiatric symptoms. *World Psychiatry*, *11*(1), 16–21. doi:10.1016/j.wpsyc.2012.01.004

Westen, D., Betan, E., & DeFife, J. A. (2011). Identity disturbance in adolescents: Associations with borderline personality disorder. *Development and Psychopathology*, *23*, 305–313. doi:10.1017/S0954579410000817

Westen, D., Dutra, L., & Shedler, J. (2005). Assessing adolescent personality pathology: Quantifying clinical judgment. *British Journal of Psychiatry*, *186*, 227–238. doi:10.1192/bjp.186.3.227

Westen, D., Gabbard, G. O., & 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, NY: Guilford Press.

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. doi:10.1521/pedi.17.4.351.23967

Westen, D., & Muderrisoglu, S. (2006). Clinical assessment of pathological personality traits. *American Journal of Psychiatry*, *163*, 1285–1287. doi:10.1176/appi.ajp.163.7.1285

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. doi:10.1037/0022-006X.65.3.429

Westen, D., Nakash, O., Thomas, C., & Bradley, R. (2006). Clinical assessment of attachment patterns and personality disorder in adolescents and adults. *Journal of Consulting and Clinical Psychology*, *74*, 1065–1085. doi:10.1037/0022-006X.74.6.1065

Westen, D., & Shedler, J. (1999a). Revising and assessing Axis II, Part 1: 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, Part 2: 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. doi:10.1521/pedi.2000.14.2.109

Westen, D., & Shedler, J. (2007). Personality diagnosis with the Shedler–Westen Assessment Procedure (SWAP): Integrating clinical and statistical measurement and prediction. *Journal of Abnormal Psychology*, *116*, 810–822. doi:10.1037/0021-843X.116.4.810

Westen, D., Shedler, J., & Bradley, R. (2006). A prototype approach to personality disorder diagnosis. *American Journal of Psychiatry*, *163*, 846–856. doi:10.1176/appi.ajp.163.5.846

Westen, D., Shedler, J., Bradley, B., & DeFife, J. A. (2012). An empirically-derived taxonomy for personality diagnosis: Bridging science and practice in conceptualizing personality. *American Journal of Psychiatry*, *169*, 273–284. doi:10.1176/appi.ajp.2011.11020274

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. doi:10.1176/appi.ajp.160.5.952

Westen, D., Waller, N., Shedler, J., & Blagov, P. S. (in press). Dimensions of personality and personality pathology: Factor structure of the Shedler–Westen Assessment Procedure-II (SWAP-II). *Journal of Personality Disorders*. doi: 10.1521/pedi_2012_26_059.

Widiger, T. A. (2005). A dimensional model of personality disorder. *Current Opinion in Psychiatry*, *18*(1), 41–43.

Widiger, T. A., & Samuel, D. B. (2005). Evidence-based assessment of personality disorders. *Psychological Assessment*, *17*, 278–287. doi:10.1037/1040-3590.17.3.278

Zimmerman, M., Chelminski, I., Young, D., Dalrymple, K., & Martinez, J. (2011). Does DSM-IV already capture the dimensional nature of personality disorders? *Journal of Clinical Psychiatry*, *72*, 1333–1339. doi:10.4088/JCP.11 m06974

Zimmerman, M., Rothschild, L., & Chelminski, I. (2005). The prevalence of DSM-IV personality disorders in psychiatric outpatients. *American Journal of Psychiatry*, *162*, 1911–1918. doi:10.1176/appi.ajp.162.10.1911

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix S1. Personality disorder scales, coefficient alphas, and corrected item–scale correlations from the derivation sample ($n = 481$).