

Personality Diagnosis with the Shedler-Westen Assessment Procedure (SWAP): Bridging the Gulf Between Science and Practice

Jonathan Shedler, PhD,¹ Drew Westen, PhD²

Editor's Note: Like the immediately preceding chapter by Blatt and his collaborators, this paper presents another of the major measures (also indicated in the review by Wallerstein in his chapter on psychoanalytic therapy research, p. 511) created to empirically delineate personality attributes and character types. The predominant focus is on the use of the SWAP for character diagnosis, with direct application as an outcome measure for underlying personality (or structural) change in psychotherapy.

The authors begin with a sophisticated discussion of the limitations of the DSM system for diagnosing personality disorders (Axis II) with respect to both diagnosis and treatment planning, before describing the development of their own diagnostic approach, the SWAP. Rather than trying to exclude clinical judgment from psychological assessment, the SWAP harnesses and systematizes clinical judgment in a manner that enhances, rather than diminishes, reliability and validity. It is an empirical approach that can bridge the longstanding schism between the experience of clinical practice and the findings of incrementally accruing science. Like Blatt et al.'s Object Relations Inventory (ORI), the SWAP can be used to test the efficacy and the effectiveness of psychoanalytically-based (as well as other) psychotherapies.

It is often easier to hear criticism when it is directed at others rather than ourselves. In this spirit, we offer an account of an unfortunate experience. During a routine medical exam, a friend had an abnormal finding on a lab test. His physician ordered more tests, then referred him to an oncologist. The oncologist ordered more tests, then referred him to a team of oncology specialists, researchers at the cutting edge of their discipline. My friend underwent a liver biopsy. The oncologists diagnosed advanced liver cancer and told him he had only months to live.

¹ Associate Professor of Psychology, University of Denver; Clinical Professor of Psychiatry, University of Colorado Health Sciences

² Professor, Department of Psychology and Department of Psychiatry and Behavioral Sciences, Emory University

In the ensuing panic there were few voices of reason. One happened to be that of a psychoanalyst, my friend's senior colleague. She asked a simple question: Had he been *feeling* sick? He had not. In fact, he had been feeling energetic and strong. The psychoanalyst raised an eyebrow. Her wordless gesture spoke volumes: Something did not add up. The pieces did not fit. If my friend had advanced liver cancer, he would likely be deathly ill.

Indeed, he did not have cancer. After additional biopsies (and ineffable emotional turmoil), the oncologists eventually concluded that his liver had an area of dense blood vessel growth (hemangioma) that had probably been present from birth and was of no medical consequence. One might reasonably ask how these research-oriented oncologists had gotten it so wrong and why an elderly psychiatrist who had not practiced medicine in decades had shown greater diagnostic acumen. No doubt many factors were at work, but we believe one factor is that the oncologists focused on laboratory findings to the exclusion of other meaningful data, including the data afforded by their own eyes and ears. Additionally, they failed to consider how the data *fit together*. Had the laboratory findings been contextualized by what else the doctors knew or could have known about their patient, they may have regarded them differently—as pieces of a diagnostic puzzle, not the diagnostic picture in its entirety. To the extent that they relied on laboratory technology to the exclusion of clinical observation, judgment, and inference, the oncologists functioned more as *technicians* than as *clinicians*.

In recent decades, the mental health professions have also emphasized data from the research laboratory over data from the clinical consulting room. Personality diagnosis once depended upon expert clinical judgment and inference about subtle, textured, and nuanced personality processes. Clinicians considered a range of data, relying not just on what patients said, but also on how they said it, drawing complexly determined inferences from patients' accounts of their lives and important relationships, from their manner of interacting with the clinician, and from their own emotional reactions to the patient. For example, expert clinicians tend not to assess lack of empathy, a diagnostic criterion for narcissistic personality disorder, by administering self-report questionnaires or asking direct questions. Often, an initial sign of lack of empathy on the part of the patient is a subtle sense on the part of the clinician of being interchangeable or replaceable, of being treated as a sounding board rather than as a fellow human being. The clinician might go on to consider whether she consistently feels this way with this particular patient, and whether such feelings are characteristic for her in her role as therapist. She might then become aware that the patient's descriptions of important others come across as somewhat two-dimensional, or that he tends to describe others more in terms of the functions they serve or the needs they meet than in terms of who they are as people. The clinician might further consider whether and how these issues dovetail with the facts the patient has provided about his life, with the problems that brought him to treatment, with

information gleaned from family members or other collateral contacts, and so on. When clinicians function as clinicians and not as technicians, it is this kind of thinking, reasoning, and inference that they engage in. Such clinical inference lies at the heart of psychodynamic approaches to understanding personality.

It is just such clinical judgment and inference that psychiatry and psychology have turned away from. As successive editions of the DSM have minimized the role of clinical judgment and inference, personality diagnosis has evolved into a largely technical task of tabulating behavioral signs and symptoms with relatively little consideration for how they fit together, the psychological functions they serve, their meanings, the developmental trajectory that gave rise to them, or the present-day factors that serve to maintain them. Indeed, the diagnostic “gold standard” in personality disorder research is the structured research interview. Such assessment methods are designed to achieve interrater reliability by minimizing the role of clinical judgment or reducing it to the lowest common denominator. Instead of relying on clinical knowledge, complexly determined inferences, and consistent impressions made on the harnessed subjectivities of seasoned therapists (McWilliams, 1999), such assessment procedures substitute standardized questions and decision rules. Indeed, they are typically not administered by expert clinicians at all, but by research assistants or trainees. Like the oncologists in the story, practitioners who rely on such diagnostic methods are functioning more as technicians than as clinicians.

We must keep in mind, however, that the DSM, and the structured assessment instruments it spawned, developed in the directions they have for good reason. Prior to DSM-III, psychiatric diagnosis was unsystematic, overly subjective, and of questionable scientific merit. It often revealed more about the clinician’s background and theoretical commitments than it did about the patient. The DSM and structured personality assessment methods evolved in the service of science, and in reaction against the unsystematic and overly subjective diagnostic methods of the past. In the evolution of diagnosis from a largely subjective, clinical enterprise to a largely technical, research-driven enterprise, much has been gained just as much has been lost. The solution cannot be to turn back the clock and abandon the technical developments of the past decades, any more than it would make sense for oncologists to disregard the most technically advanced laboratory tests available. The solution, rather, may be a marriage of the best aspects of clinical wisdom and empirical rigor. We need not choose between empiricism devoid of clinical realism versus a return to a pre-empirical past. To borrow the paradoxical title of a popular movie, progress may lie in going back to the future

This chapter describes the Shedler-Westen Assessment Procedure (SWAP), an approach to personality assessment designed to *harness* clinical judgment and inference rather than eliminate it, and combine the best features of the clinical

and empirical traditions in personality assessment. It renders clinical constructs accessible to empirical investigation and provides a means of assessing personality that is both dynamically relevant and empirically grounded.

This chapter will (1) review problems inherent in the current DSM diagnostic system, (2) discuss difficulties associated with the use of clinical observation and inference in research, (3) describe the development of the SWAP and its use in systematizing clinical observation, (4) present a clinical case illustrating how the SWAP provides a bridge between descriptive psychiatry and psychodynamic case formulation, and (5) discuss recommendations for revising and refining DSM Axis II based on empirical findings from a national sample of patients.

WHY REVISE AXIS II?

The approach to the diagnosis of personality disorders codified by DSM finds little favor with either clinicians or researchers. There is consensus among researchers into personality disorders that the DSM classifications system for personality disorders requires major reconfiguration. Some of the problems with DSM-Axis II include the following (see Westen & Shedler, 1999a, 2000):

1. The diagnostic categories do not rest on a solid empirical foundation and often disagree with empirical findings from cluster and factor analyses (Blais & Norman, 1997; Clark, 1992; Harkness, 1992; Livesley & Jackson, 1992; Morey, 1988).
2. DSM-Axis II artificially dichotomizes continuous variables (diagnostic criteria) into present/absent, which is neither theoretically nor statistically sensible (e.g., how little empathy constitutes lack of empathy?).
3. DSM-Axis II commits arbitrarily to a categorical diagnostic system. It may be more useful to conceptualize borderline personality pathology, for example, on a continuum from none through moderate to severe, rather than classifying borderline personality disorder as present/absent (Widiger, 1993).
4. DSM-Axis II lacks the capacity to weight criteria that differ in their diagnostic importance (Davis, Blashfield & McElroy, 1993).
5. Comorbidity between diagnoses of different personality disorders is unacceptably high. Patients who receive any personality disorder diagnosis often receive four to six out of a possible ten (Blais & Norman, 1997, Grilo, Sanislow & McGlashan, 2002, Oldham, Skodol, Kellman, Hyler, et al., 1992, Pilkonis, Heape, Proietti, et al., 1995, Watson & Sinha, 1998), indicating lack of discriminant validity of the diagnostic constructs, the assessment methods, or both.
6. In trying to reduce comorbidity, DSM work groups have had to gerrymander diagnostic categories and criteria, sometimes in ways faithful

neither to clinical observation nor empirical data. For example, they excluded lack of empathy and grandiosity from the diagnostic criteria for antisocial personality disorder to minimize comorbidity with narcissistic personality disorder, despite evidence that these traits are associated with both disorders (Widiger & Corbitt, 1995).

7. Efforts to define personality disorders more precisely have led to narrower criterion sets over time, progressively eroding the distinction between personality *disorders* (multifaceted syndromes encompassing cognition, affectivity, motivation, interpersonal functioning, and so on) and simple personality *traits*. The diagnostic criteria for paranoid personality disorder, for example, are essentially redundant indicators of a single trait, chronic suspiciousness. The diagnostic criteria no longer describe the multifaceted personality syndrome recognized by most clinical theorists (Millon, 1990; Millon & Davis, 1997).
8. DSM-Axis II fails to consider personality strengths that might rule out personality disorder diagnoses for some patients. For example, differentiating between a patient with narcissistic personality disorder and a much healthier person with narcissistic dynamics may not be a matter of counting symptoms, but rather of noting whether the patient has such positive qualities as the capacity to love and to sustain meaningful relationships characterized by mutual caring and sharing.
9. DSM-Axis II does not encompass the spectrum of personality pathology that clinicians see in practice. Among patients receiving treatment for personality pathology (defined as enduring, maladaptive patterns of emotion, thought, motivation, or behavior that lead to distress or dysfunction), fewer than 40% can be diagnosed on axis II (Westen & Arkowitz-Westen, 1998).
10. The categories and criteria are not as clinically useful or relevant as they might be. For example, knowing whether a person meets DSM-IV diagnostic criteria for avoidant personality disorder or dependent personality disorder tells us little about the meaning of the person's symptoms, which personality processes to target for treatment, or how to treat them.
11. The algorithm used for diagnostic decisions (counting symptoms) diverges from the methods clinicians use, or could plausibly be expected to use, in real-world practice. Research in cognitive science indicates that clinicians are unlikely to make diagnoses by additively tabulating symptoms. Rather, they gauge the overall "match" between a patient and a mental template or prototype of the disorder (i.e., they consider the features of a disorder as a configuration or gestalt), or they use causal theories to make sense of the functional relations between the symptoms

(Blashfield, 1985; Cantor & Genero, 1986; Kim & Ahn, 2002; Westen, Heim, Morrison, Patterson, & Campbell, 2002).

12. The instruments used to assess personality disorders do not meet the standards for validity normally expected in psychological research (Perry, 1992; Skodol, Oldham, Rosnick, Kellman, & Hyler, 1991; Westen, 1997) and they show poor test-retest reliability at intervals greater than 6 weeks (First, Spitzer, Gibbon, Williams, et al., 1995; Zimmerman, 1994). The lack of test-retest reliability is especially problematic, given that personality disorders are, by definition, enduring and stable over time.³

Most of the proposed solutions to these problems share the assumption that progress lies in further minimizing the role of the clinician, either by developing increasingly behavioral and less inferential diagnostic criteria, or by bypassing clinical judgment entirely through self-report questionnaires. Such attempted solutions may, however, be part of the problem. By eliminating clinical observation and inference, we may unintentionally be eliminating from study the psychological phenomena that are of greatest relevance and importance. The empathically attuned clinician may still be the only “measurement instrument” sensitive enough to register crucial psychological phenomena (Shedler, Mayman, & Manis, 1993). An alternative to trying to eliminate clinical observation and inference is to *harness* it for scientific use.

THE PROBLEM WITH CLINICAL DATA

The problem with clinical observation and inference is *not* that it is inherently unreliable, as some researchers seem to assume. The problem is that it tends to come in a form that does not lend itself readily to systematic study (Shedler, 2002). Rulers measure in inches and scales measure in pounds, but what metric do psychotherapists share? Imagine three clinicians reviewing the same case material. One might speak of schemas and belief systems, another of conditioning history, and the third of transference and resistance. Even among psychoanalytic practitioners, there is little consensus about constructs and terminology. One psychoanalyst might speak of conflict and compromise formation, a second of object relations, and the third, perhaps, of self defects.

It is not readily apparent whether the hypothetical clinicians can or cannot make similar observations and inferences. There are three possibilities: (1) They may be observing and describing the same thing but using different language and metaphor systems to express it. (2) They may simply be attending to differ-

³ Poor test-retest reliability has led some researchers to suggest that personality disorders may be less stable than previously believed. Another hypothesis is that the assessment instruments overemphasize transitory behavioral symptoms (e.g., self-cutting in borderline patients) and underemphasize underlying personality processes that are, in fact, more stable over time (e.g., emotional dysregulation and self-hatred in borderline patients).

ent aspects of the clinical material, as in the parable of the elephant and the blind men. (3) They may not be able to make the same clinical observations at all. *If we want to know whether clinicians can make the same observations and inferences, we must ensure that they speak the same language and pay attention to the same range of clinical phenomena.*

A STANDARD VOCABULARY FOR CASE DESCRIPTION

The *Shedler-Westen Assessment Procedure* (SWAP) is an assessment instrument designed to provide clinicians of all theoretical orientations with a standard vocabulary with which to express their observations and inferences about personality functioning (Shedler & Westen, 1998; 2004a, 2004b; Westen & Shedler, 1999a, 1999b). The vocabulary consists of 200 statements, each printed on a separate index card. Each statement may describe a given patient very well, somewhat, or not at all. A clinician who knows a patient well can describe the patient by ranking or ordering the statements into 8 categories, from those that are most descriptive of the patient (assigned a value of 7) to those that are not descriptive (assigned a value of 0). Thus, the SWAP yields a score from 0 to 7 for each of 200 personality-descriptive variables. An interactive, web-based version of the SWAP can be previewed at www.SWAPassessment.com.

The standard vocabulary of the SWAP allows clinicians to provide in-depth psychological descriptions of patients in a systematic and quantifiable form. It also ensures that all clinicians attend to the same spectrum of clinical phenomena. SWAP statements are written in a manner close to the data (e.g., “Tends to be passive and unassertive,” or “Emotions tend to spiral out of control, leading to extremes of anxiety, sadness, rage, etc.”), and statements that require inference about internal processes are written in clear and unambiguous language (e.g., “Tends to blame own failures or shortcomings on other people or circumstances; attributes his/her difficulties to external factors rather than accepting responsibility for own conduct or choices”). Writing items in this jargon-free manner minimizes idiosyncratic and unreliable interpretive leaps. It also makes the item set useful for all clinicians regardless of their theoretical commitments.

The SWAP is based on the Q-Sort method, which requires clinicians to place a predetermined number of statements in each category (i.e., it uses a “fixed distribution”). The SWAP distribution resembles the right half of a normal distribution or “bell-shaped curve.” One-hundred items are placed in the “0” or not descriptive category and progressively fewer items are placed in the higher categories. Only 8 items are placed in the “7” or most descriptive category. The use of a fixed distribution has important psychometric advantages and eliminates much of the measurement error or “noise” inherent in standard rating procedures (see Block, 1978, for the psychometric rationale underlying the Q-sort method).⁴

⁴ One way it does so is by ensuring that raters are “calibrated” with one another. Consider the situation with standard rating scales, where raters can use any value as often as they wish. Inevitably, certain raters will tend toward

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

Most importantly, the SWAP-200 (the first major edition of the SWAP) is the product of a 7-year iterative revision process that incorporated the feedback of hundreds of clinician-consultants who used earlier versions of the instrument (Shedler & Westen, 1998) to describe their patients. We asked each clinician-consultant one crucial question: "Were you able to describe the things you consider psychologically important about your patient?" We added, rewrote, and revised items based on this feedback, then asked new clinician-consultants to describe new patients. We repeated this process over many iterations until most clinicians could answer "yes" most of the time. A newer, revised version of the SWAP item set, the SWAP-II incorporates the additional feedback of over 2000 clinicians of all theoretical orientations. The iterative item revision process was designed to ensure both the comprehensiveness and the clinical relevance of the SWAP item sets.

Because the SWAP is jargon-free and clinically comprehensive, it has the potential to serve as a universal language for describing personality pathology. Our studies demonstrate that experienced clinicians of diverse theoretical orientations understand the items and can apply them reliably to their patients. In one study, a nationwide sample of 797 experienced psychologists and psychiatrists of diverse theoretical orientations, who had an average of 18 years practice experience post training, used the SWAP-200 to describe patients with personality pathology (Westen and Shedler, 1999a). These experienced therapists provided similar SWAP-200 descriptions of personality disorders regardless of their theoretical commitments, and fully 72.7% agreed with the statement, "I was able to

extreme values (e.g., values of 0 and 7 on a 0-7 scale) whereas others will tend toward middle values (e.g., values of 4 and 5). Thus, the ratings reflect not only the characteristics of the patients but also the calibration of the raters. The Q-Sort method, with its fixed distribution, eliminates this kind of measurement error, because all clinicians must use each value the same number of times. If use of a standard item set gives clinicians a common vocabulary, use of a fixed distribution can be said to give them a common "grammar" (Block, 1978).

express most of the things I consider important about this patient” (the highest rating category). In a subsequent sample of 1201 psychologists and psychiatrists who used the SWAP-II, over 80% “agreed” or “strongly agreed” with the statement, “The SWAP-II allowed me to express the things I consider important about my patient’s personality” (less than 5% disagreed). These ratings were unrelated to clinicians’ theoretical orientation. Virtually identical agreement rates were obtained in a national sample of clinicians who used the adolescent version of the instrument, the SWAP-II-A.

PSYCHODYNAMICS WITHOUT JARGON

Some investigators have assumed that clinical concepts, especially psychodynamic constructs, are too vague, theoretical, or hypothetical to study empirically. The following SWAP-II items illustrate how the instrument operationalizes some psychodynamic concepts (focusing, for purposes of illustration, on defensive processes). Note that the constructs—rinsed of theoretical jargon—are relevant to a wide range of clinicians, irrespective of theoretical commitments. The traditional psychoanalytic terms for the concepts (which are not included in the SWAP) are in brackets:

SWAP item #	SWAP Item Text
116	Tends to see own unacceptable feelings or impulses in other people instead of in him/herself. [projection]
144	Tends to see self as logical and rational, uninfluenced by emotion; prefers to operate as if emotions were irrelevant or inconsequential. [intellectualization]
78	Tends to express anger in passive and indirect ways (e.g., may make mistakes, procrastinate, forget, become sulky, etc.). [passive aggression]
14	Tends to blame own failures or shortcomings on other people or circumstances; attributes his/her difficulties to external factors rather than accepting responsibility for own conduct or choices. [externalization]
45	Is prone to idealizing people; may see admired others as perfect, larger than life, all wise, etc. [idealization]
165	Tends to distort unacceptable wishes or feelings by transforming them into their opposite (e.g., may express excessive concern while showing signs of unacknowledged hostility, disgust about sexual matters while showing signs of unacknowledged excitement, etc.). [reaction formation]

AN ILLUSTRATION: BORDERLINE PERSONALITY PATHOLOGY

Some clinicians may doubt that a finite set of 200 statements can capture the richness and complexity of clinical case description. However, SWAP statements can be combined in virtually infinite patterns to express subtle clinical concepts. The mathematically inclined reader might note that there are 200 factorial possible orderings of the SWAP statements (an inexpressibly large number). The musically inclined reader might note that all of Western music can be notated using combinations of only 12 tones.

Many clinical theorists consider *splitting*, *projective identification*, and *identity disturbance* to be hallmarks of borderline personality pathology (see *Personality Patterns and Disorders*, p. 15). Consider, for example, the items reproduced below from the original SWAP-200 item set. The three items, *taken in combination*, convey something of the defensive splitting seen in patients with borderline personality pathology:

SWAP item #	SWAP Item Text
162	Expresses contradictory feelings or beliefs without being disturbed by the inconsistency; has little need to reconcile or resolve contradictory ideas.
45	Tends to idealize certain others in unrealistic ways; sees them as “all good,” to the exclusion of commonplace human defects.
79	Tends to see certain others as “all bad,” and loses the capacity to perceive any positive qualities the person may have.

The following items, from the SWAP-II, capture some additional meanings of the concept *splitting*.

SWAP item #	SWAP Item Text
9	When upset, has trouble perceiving both positive and negative qualities in the same person at the same time (e.g., may see others in black or white terms, shift suddenly from seeing someone as caring to seeing him/her as malevolent and intentionally hurtful, etc.).
18	Tends to stir up conflict or animosity between other people (e.g., may portray a situation differently to different people, leading them to form contradictory views or work at cross purposes).

The next group of items, *taken in combination*, captures at least one meaning of the term *projective identification*:

SWAP item #	SWAP Item Text
116	Tends to see own unacceptable feelings or impulses in other people instead of in him/herself.
76	Manages to elicit in others feelings similar to those s/he is experiencing (e.g., when angry, acts in such a way as to provoke anger in others; when anxious, acts in such a way as to induce anxiety in others).
154	Tends to draw others into scenarios, or “pull” them into roles, that feel alien or unfamiliar (e.g., being uncharacteristically insensitive or cruel, feeling like the only person in the world who can help, etc.).

The concept *identity disturbance* (or *identity diffusion*) subsumes a wide range of phenomena (Wilkinson-Ryan & Westen, 2000). When the same term has been used in the literature in different ways, or used differently by different theorists, we wrote multiple SWAP items to cover the multiple meanings. The following SWAP-II items illustrate some of the manifestations and facets of identity disturbance:

SWAP item #	SWAP Item Text
15	Lacks a stable sense of who s/he is (e.g., attitudes, values, goals, and feelings about self seem unstable or ever-changing).
151	Appears to experience the past as a series of disjointed or disconnected events; has difficulty giving a coherent account of his/her life story.
90	Is prone to painful feelings of emptiness (e.g., may feel lost, bereft, abjectly alone even in the presence of others, etc.).
172	Seems unable to settle into, or sustain commitment to, identity-defining life roles (e.g., career, occupation, lifestyle, etc.).
150	Tends to identify with admired others to an exaggerated degree, taking on their attitudes, mannerisms, etc. (e.g., may be drawn into the “orbit” of a strong or charismatic personality).
87	Sense of identity revolves around a “cause,” movement, or label (e.g., adult child of alcoholic, adult survivor, environmentalist, born-again Christian, etc.); may be drawn to extreme or all-encompassing belief systems.
38	Tends to feel s/he is not his/her true self with others; may feel false or fraudulent.
102	Has a deep sense of inner badness; sees self as damaged, evil, or rotten to the core (whether consciously or unconsciously).

The next group of items helps flesh out a picture of a certain kind of borderline patient, addressing issues of affect regulation, object relations, cognition, and so on:

SWAP item #	SWAP Item Text
191	Emotions tend to change rapidly and unpredictably.
12	Emotions tend to spiral out of control, leading to extremes of anxiety, sadness, rage, etc.
185	Is prone to intense anger, out of proportion to the situation at hand (e.g., has rage episodes).
157	Tends to become irrational when strong emotions are stirred up; may show a significant decline from customary level of functioning.
117	Is unable to soothe or comfort him/herself without the help of another person (i.e., has difficulty regulating own emotions).
98	Tends to fear s/he will be rejected or abandoned.
11	Tends to become attached quickly or intensely; develops feelings, expectations, etc. that are not warranted by the history or context of the relationship.
167	Is simultaneously needy of, and rejecting toward, others (e.g., craves intimacy and caring, but tends to reject it when offered).
153	Relationships tend to be unstable, chaotic, and rapidly changing.
52	Has little empathy; seems unable or unwilling to understand or respond to others' needs or feelings.
176	Tends to confuse own thoughts, feelings, or personality traits with those of others (e.g., may use the same words to describe him/herself and another person, believe the two share identical thoughts and feelings, etc.).
41	Appears unable to describe important others in a way that conveys a sense of who they are as people; descriptions of others come across as two-dimensional and lacking in richness.
29	Has difficulty making sense of other people's behavior; tends to misunderstand, misinterpret, or be confused by others' actions and reactions.

The last group of items includes descriptors that might apply to a more disturbed borderline patient, perhaps one likely to be seen in an inpatient setting (Gunderson, 2001):

SWAP item #	SWAP Item Text
134	Tends to act impulsively (e.g., acts without forethought or concern for consequences).
142	Tends to make repeated suicidal threats or gestures, either as a “cry for help” or as an effort to manipulate others.
109	Tends to engage in self-mutilating behavior (e.g., self-cutting, self-burning, etc.).
188	Work-life and/or living arrangements tend to be chaotic or unstable (e.g., job or housing situation seems always temporary, transitional, or ill-defined).
44	When distressed, perception of reality can become grossly impaired (e.g., thinking may seem delusional).

The items reproduced here are illustrative only and are not intended to describe “the” borderline patient or any particular borderline patient. They are intended only to illustrate that it is possible to describe clinically and psychodynamically important constructs without succumbing to either reductionism or jargon. Further, such descriptions are empirically testable.

A Case Illustration

At present, descriptive psychiatric diagnosis and clinical case formulation are largely independent activities. The former is aimed at classification (a nomothetic approach) whereas the latter is aimed at understanding an individual patient (an idiographic approach). The SWAP bridges these activities. It provides a dimensional diagnosis score for each personality disorder included in DSM-IV (as well as for the additional personality disorders proposed in the DSM-IV appendix). It also provides a richly detailed clinical case narrative relevant to case formulation and treatment planning.

Dimensional personality disorder (PD) scores measure the “fit” or “match” between an individual patient and prototype descriptions representing each personality disorder in its “ideal” or pure form (e.g., a prototype description of paranoid personality disorder). Thus, each personality disorder is diagnosed on a continuum; a low PD score indicates that the patient does not fit or match the personality disorder syndrome and a high PD score indicates that the patient matches it well (with intermediate scores indicating varying degrees of “fit”). PD scores can be graphed to create a personality disorder profile resembling an MMPI profile, as illustrated in Figure 1. Dimensional diagnosis is consistent with clinical thinking and advocated by virtually all contemporary personality researchers (Widiger & Simonsen, 2005). (We are developing web-based software that will

allow clinicians to input SWAP scores and receive a computer-generated interpretive report with personality disorder diagnoses, a detailed case description, and treatment recommendations. Visit *www.SWAPassessment.com* to preview the software,).

A clinical case example may best illustrate these diagnostic applications of the SWAP.⁵

Case Background

“Melania” is a 30 year old Caucasian woman. Her presenting complaints included substance abuse and inability to extricate herself from an abusive relationship. The initial assessment included a psychiatric intake interview and administration of both the SCID and SCID-II structured interviews. She met SCID criteria for an Axis I diagnosis of substance abuse and SCID-II criteria for an Axis II diagnosis of borderline personality disorder with histrionic traits. The intake interviewer assigned a score of 45 on the Global Assessment of Functioning (GAF) scale, indicating severe symptoms and impairment in functioning.

Melania’s early family environment was marked by neglect and parental strife. A recurring family scenario is illustrative: Melania’s mother would scream at her husband, telling him he was a failure and that she was going to leave him; she would then slam the door and lock herself in her room, leaving Melania frightened and in tears. Both parents would then ignore Melania, often forgetting to feed her. Melania’s parents divorced when she was 8. After the divorce, Melania lived with her mother, who showed little concern for her needs or welfare.

By adolescence, Melania had developed behavioral problems. She often skipped school and spent her days sleeping or wandering the streets. At age 18, she left home and began what she described as “life on the streets.” She engaged in a series of impulsive, chaotic, and rapidly changing sexual relationships which led to three abortions by age 24. She abused street drugs, eventually developing a pattern of cocaine and heroin abuse (snorting). She also engaged in petty criminal activity, including shoplifting and stealing from employers.

Melania held a series of low paying jobs that were not commensurate with her ability or education. She failed to hold any job for more than a few months and was fired from each when she was caught stealing. In her mid-twenties, Melania moved in with her boyfriend, a small-time drug dealer who exploited her financially and abused her physically. He spent his days sleeping or watching television while Melania worked to pay the rent. She often had sex with other men

⁵ The material presented in this section is adapted from Lingiardi, Shedler, & Gazzillo (2006).

to obtain money or drugs for her boyfriend. He sometimes beat her when he was dissatisfied with what she brought home.

Melania began psychodynamic therapy at a frequency of three sessions per week. The first ten psychotherapy sessions were tape recorded and transcribed. Two clinicians (blind to all other data) reviewed the transcripts and provided SWAP-200 descriptions of Melania, based on the information contained in the session transcripts. The SWAP-200 scores were then averaged across the two clinical judges to obtain a single SWAP-200 description.⁶ After two years of psychotherapy, ten consecutive psychotherapy sessions were again recorded and transcribed, and the SWAP assessment procedure was repeated.

Personality Disorder Diagnosis

The solid line in Figure 1 shows Melania's PD scores at the beginning of treatment for the ten personality disorders included in DSM-IV. A "healthy functioning" index is graphed as well, which reflects clinicians' consensual understanding of healthy personality functioning (Westen & Shedler, 1999a). For ease of interpretation, the PD scores have been converted to T-scores (Mean = 50, SD = 10) based on norms established in a psychiatric sample of patients with axis II diagnoses (Westen & Shedler, 1999a). Although the SWAP assesses personality disorders dimensionally and treats each personality disorder diagnosis as a continuum, we have established cutoff scores for "backward compatibility" with DSM-IV. To maintain continuity with the DSM-IV categorical diagnostic system, we have suggested T = 60 as a threshold for making a categorical personality disorder diagnosis, and T = 55 as a threshold for diagnosing "features."⁷

Melania's PD profile shows a marked elevation for borderline personality disorder (T = 65.4, approximately one and a half standard deviations above the sample mean), with secondary elevations for histrionic personality disorder (T = 56.6) and antisocial personality disorder (T = 55.7). Applying the recommended cutoff scores, her axis II diagnosis is borderline personality disorder with histrionic and antisocial features. Also noteworthy is the T-Score of 41 for the "healthy functioning" index, nearly a standard deviation below the mean in a reference sample of patients with Axis II diagnoses. The low score indicates significant impairment in functioning and parallels the low GAF score assigned by the intake interviewer.

⁶ Averaging across raters enhances the reliability of the resulting scores.

⁷ The relatively low thresholds reflect the fact that the reference sample consisted of patients with a diagnosis of personality disorder. Thus, a T-score of 50 indicates "average" functioning among patients with personality disorder diagnoses, and a T-score of 60 represents an elevation of one standard deviation relative to other patients with personality disorder diagnoses.

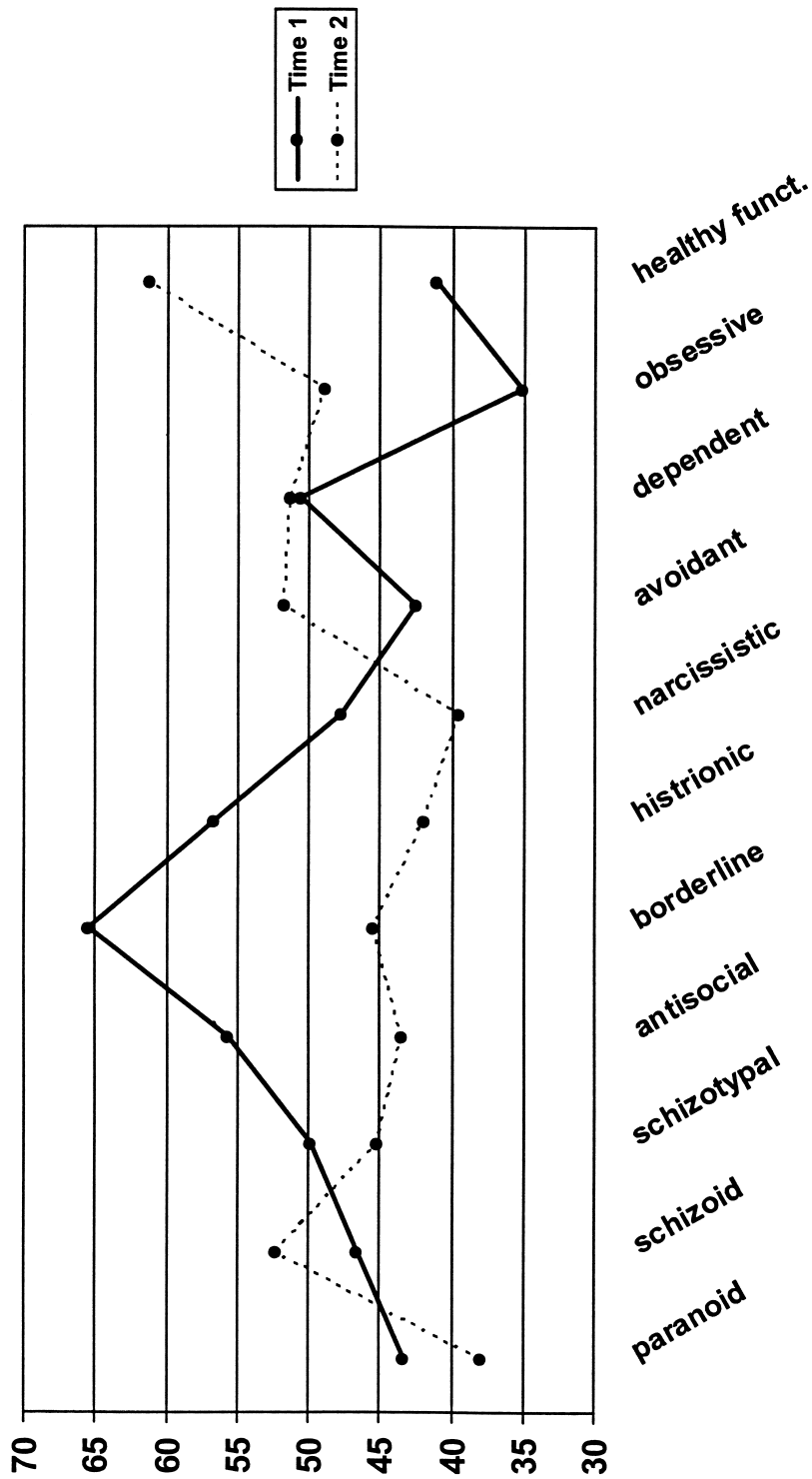


Figure 1: PD score profile

Narrative Case Description

We can generate a narrative case description by listing the SWAP items assigned the highest scores in the patient's SWAP description (e.g., items with scores of 5, 6, and 7). Below is a narrative case description for "Melania" based on the top 30 most descriptive SWAP-200 items. We have grouped together conceptually related items. To aid the flow of the text, we have made some minor grammatical changes and added connecting text. However, the SWAP-200 items are reproduced essentially verbatim. The narrative description is based on the same data used to generate the PD score profile in Figure 1.

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

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

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

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

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

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

Assessing Change in Therapy

The case of Melania has a happy ending. After two years of psychotherapy, the SWAP assessment revealed significant personality changes. The changes parallel concrete behavior changes as well as changes in Melania's life circumstances (e.g., ending her drug abuse, getting and keeping a good job, ending her involvement with her abusive boyfriend, no longer engaging in theft, promiscuous sex, or prostitution).

The dotted line in Figure 1 shows Melania's PD scores after two years of psychotherapy. Her scores on the borderline, histrionic, and antisocial dimensions have dropped below $T = 50$ and she no longer warrants a diagnosis of personality disorder. Her score on the healthy functioning index has increased by two standard deviations, from 41.0 to 61.2.

To assess change in an ideographic, more fine-grained manner, we created a change score for each individual SWAP item by subtracting the item score at Time 1 from the score at Time 2. The narrative description of change, below, is comprised of the SWAP items with change scores > 4 . Again, we have made some minor grammatical changes and added connecting text to aid the flow of the text, but the SWAP-200 items are reproduced essentially verbatim.

Melania has developed strengths and inner resources that were not evident at the Time 1 assessment. She has come to terms with painful experiences from the past, finding meaning in, and growing from, these experiences; she has become more articulate and better able to express herself in words; she has a newfound ability to appreciate and respond to

humor; she is more capable of recognizing alternative viewpoints, even in matters that stir up strong feelings; she is more empathic and sensitive to other's needs and feelings; and she is more likeable.

There is marked improvement in many areas associated specifically with borderline psychopathology. With respect to affect regulation, Melania is less prone to become irrational when strong emotions are stirred up, is more likely to express affect appropriate in quality and intensity to the situation at hand, and is better able to sooth or comfort herself when distressed. She is less prone to confuse her own thoughts and feelings with those of others, less manipulative, and less likely to devalue others and see them as "all bad." She has come to terms with negative feelings toward her parents.

Melania is also less impulsive, more conscientious and responsible, and more aware of the consequences of her actions. Her living arrangements are more stable, as is her work life. Melania's use of illicit drugs has decreased significantly, and she is no longer drawn to abusive relationships.

As the more severe aspects of borderline personality pathology have receded, other conflicts and symptoms have moved to the fore. For example, Melania appears to have developed somewhat obsessional defenses against painful affect. She adheres more rigidly to daily routines and becomes anxious or uncomfortable when they are altered. She is more prone to think in an abstract and intellectualized manner, and tries to see herself as more logical and rational, less influenced by emotion.

Despite her wish to act more logically and rationally, Melania seems engaged in an active struggle to control her affect and impulses. She tends to oscillate between undercontrol and overcontrol of needs and wishes, either expressing them impulsively or disavowing them entirely. She has more difficulty allowing herself to experience strong pleasurable emotions (e.g., excitement, joy). She is more prone to repress, "forget," or otherwise distort distressing events.

Finally, there are changes in Melania's relationships and orientation toward sexuality. Whereas before she presented in a histrionic manner (i.e., with exaggerated feminine traits), she is now more disparaging of traditionally feminine traits, instead emphasizing independence and achievement. Whereas previously she engaged in multiple chaotic sexual relationships, she now seems conflicted about her intimacy needs. She craves intimacy but tends to reject it when offered. She has more difficulty directing both sexual and tender feelings toward the same person, seeing men as either respectable and virtuous, or sexy and exciting, but not both. She is more likely to hold grudges.

We leave it to readers to judge the clinical relevance of the SWAP and the value of the diagnostic profiles and narrative case descriptions that it yields. Note, however, that the standard vocabulary of the SWAP ensures that different clinicians will describe the same patient in much the same way, once they learn to use the SWAP reliably. Had other clinicians described Melania using the SWAP, the narrative descriptions would have been much the same, since the descriptive statements comprising the narrative were taken directly from the SWAP-200 item set.

RELIABILITY AND VALIDITY

Researchers in psychology and psychiatry have often assumed that clinical observation and judgment is unreliable, and a well-established literature documents the limitations of “clinical judgment.” Unfortunately, studies of “clinical judgment” have often asked clinicians to make judgments about things that fall well outside their legitimate area of expertise (equally unfortunate, some clinicians have been all too willing to offer such prognostications). More problematic, the studies have typically conflated clinicians’ ability to make accurate observations and inferences (which they do well) with their ability to combine and weight variables to derive optimal predictions (a task that is *necessarily* performed better by statistical methods such as regression equations). In fact, a substantial literature documents the reliability and validity of clinical observation and inference *when it is quantified and used appropriately* (see Westen & Weinberger, 2004, for a detailed discussion and literature review).

The SWAP differs from past approaches in that it harnesses clinical judgment using psychometric methods developed specifically for this purpose, then applies statistical and actuarial methods to the resulting variables. In short, it relies on clinicians to do what they do best: making specific behavioral observations and inferences about the individual patients they know well. It relies on statistical algorithms to do what they do best: combining data optimally to derive reliable and valid diagnostic scales and indices.

Inter-rater reliability of SWAP-200 PD scale scores (Figure 1) and other diagnostic scales derived from the SWAP is above 0.80 for all scales in all studies conducted to date (e.g., Westen and Muderrisoglu, 2003, in press) and is often above 0.90 (e.g., Marin-Avellan, McGauley, Campbell, & Fonagy, 2005). These reliability coefficients are at least as high as those typically reported for highly structured research interviews that avoid clinical inference and “just stick to the facts” (i.e., DSM-IV diagnostic criteria). Additionally, the SWAP diagnostic scales correlate highly with a wide range of external criterion measures in both adult and adolescent samples, including, e.g., genetic history variables such as psychosis in first- and second-degree relatives, substance abuse in first- and second-degree relatives,

developmental history variables such as childhood sexual and physical abuse, life events such as psychiatric hospitalizations and suicide attempts, ratings of adaptive functioning, and so on (see Shedler & Westen, 2004b; Westen & Muderrisoglu, 2003; Westen & Shedler, 1999a; Westen, Shedler, Durrett, Glass, & Martens, 2003; Westen & Weinberger, 2004).⁸

We will describe some illustrative studies in detail.⁹ Westen and Muderrisoglu (2003) interviewed a sample of outpatients using the Clinical Diagnostic Interview (CDI) (Westen, 2002), a systematic interview (2½ to 3 hours in length) designed to systematize the personality assessment methods employed by knowledgeable clinicians in real-world practice (Westen, 1997). The CDI does not ask patients to describe their own personality traits, but instead elicits narrative descriptions of patients' lives and important relationships. The narrative descriptions allow clinical interviewers to draw reliable inferences about patients' characteristic ways of thinking, feeling, regulating emotions and impulses, relating interpersonally, and so on (much as a sophisticated clinician might do in the first two to four meetings with a new patient).

The primary aims of the study were to (1) assess inter-rater reliability of SWAP diagnostic scales as assessed by independent clinicians who either conducted or observed (on videotape) the CDI interview, and (2) to assess convergent validity between these independent raters and the treating clinicians, whose SWAP scores were based on extensive contact with their patients over time. All of the clinical assessors were blind to the data provided by the others. The study examined the reliability and validity of ten SWAP PD scores plus seven other SWAP diagnostic indices (see Westen & Shedler, 1999b).

Inter-rater reliability between independent interviewers averaged greater than 0.80 for all SWAP scores. Convergent validity coefficients between interviewers and treating clinicians were also above 0.80 for all scores. Discriminant validity coefficients (i.e., correlations between unrelated diagnostic scales) were excellent, hovering near zero. To provide some reference points with which to compare these values, convergent validity between personality disorder diagnoses derived from structured research interviews and diagnoses based on the LEAD standard

⁸ Interrater reliability of the *overall* SWAP description (versus scales derived from the SWAP) is also high, ranging in prior studies from 0.75 to 0.89 (Marin-Avellan, et al., 2004; Shedler & Westen, 1998; Westen & Muderrisoglu, 2003). Reliability of the overall SWAP description is most relevant to narrative case descriptions derived from the SWAP (such as the narrative description of "Melania"). The overall reliability of a Q-sort personality description refers to agreement between raters regarding the *ordering* of the items or statements, traditionally measured by Pearson's *r* (Block, 1978). For example, a SWAP-200 description of a patient consists of one column by 200 rows of data (with each row containing the score [0 to 7] for the corresponding SWAP-200 item). Each additional rater adds one additional column of data, and interrater reliability is computed by correlating pairs of columns. Item scores are typically averaged across all available raters to obtain an aggregate Q-sort description. The reliability of the aggregate description is estimated by the Spearman-Brown formula (when there are two raters) or coefficient alpha (when there are multiple raters). The approach is identical to that used to estimate the internal reliability of a psychometric scale, except that the raters are treated as test items and the Q-sort items are treated as cases.

⁹ The material presented here is adapted from Westen and Weinberger, 2004.

(Spitzer, 1983) have ranged from 0.00 to 0.40, and discriminant validity has been quite poor (see Pilkonis, Heape, Ruddy, & Serrao, 1991; Pilkonis, et al., 1995). Similarly, a meta-analysis of personality disorder dimensions assessed via self- and informant-report yielded a median correlation of only 0.36 (Klonsky, Oltmanns, & Turkheimer, 2002).

A second study (Bradley & Westen, 2003) examined convergence between SWAP scores and patient self-report data for borderline and antisocial personality disorder (the two personality disorders for which self-report and informant-report data tend to converge). Advanced clinical psychology graduate students used the SWAP-200 to describe 54 outpatients after the fifth clinical contact hour. The patients completed the Personality Assessment Inventory (PAI) (Morey, 1991). Convergent validity was high, with SWAP antisocial and borderline personality disorder scores differentially predicting antisocial and borderline scores on the PAI. Discriminant validity coefficients were desirably low (e.g., the SWAP obsessive-compulsive PD score correlated *negatively* with PAI antisocial and borderline scores), indicating excellent diagnostic specificity. The data provide further evidence for the validity of the SWAP-200 as an assessment tool.

A study from a research group other than our own reported comparable findings (Marin-Avellan, et al., 2005). The investigators applied the SWAP-200 to audiotaped Adult Attachment Interviews (AAI) (Main, Kaplan, & Cassidy, 1985) plus chart records for a sample of inpatients at a maximum security forensic hospital (a method similar to methods for coding psychopathy; Hare, Harpur, Hakstian, Forth, et al., 1990). Interrater reliability for SWAP-200 PD scale provided by independent assessors was high, with a median interrater correlation of $r = .91$.

Additionally, SWAP PD scores differentiated patients who had committed violent versus nonviolent offenses, whereas SCID-II diagnosis did not. The SWAP-200 also proved superior to the SCID-II in predicting ward behavior, assessed independently by ward nurses (who were blind to all other data) using a 49-item interpersonal rating scale. SWAP antisocial PD scores correlated significantly with dominance behavior and coercive behavior observed on the ward, and correlated negatively with submissive behavior and compliant behavior observed on the ward. In contrast, the SCID-II predicted only dominance behavior. The findings clearly demonstrate incremental validity of the SWAP-200 relative to a widely used instrument for assessing personality disorders that relies substantially on patient self-report.

In summary, experienced clinicians can make highly reliable observations and inferences about personality dynamics, given a suitable technology for harnessing their judgments. The belief that clinicians cannot reliably assess psychodynamic and other complex clinical constructs, is mistaken.

TOWARD DSM-V: AN EMPIRICAL APPROACH TO REVISING AND REFINING DIAGNOSTIC CRITERIA

The approach to personality diagnosis codified by DSM-IV has elicited little enthusiasm from either clinicians or researchers. Ultimately, revisions to the diagnostic categories and criteria over successive editions of the DSM reflect committee decision processes, which can be influenced by group dynamics, the opinions of individual committee members, the sociopolitical *zeitgeist*, and other such factors. Here we describe an alternative, empirical approach to identifying diagnostic criteria of personality disorders.¹⁰

Identifying Core Features of Personality Disorders

Because the SWAP quantifies clinical case description, it allows investigators to statistically combine case descriptions to obtain a composite description of a particular grouping of patients. This is accomplished by averaging (aggregating) the values assigned to each SWAP item across a relevant patient sample. For example, if we obtain SWAP descriptions for a representative sample of patients diagnosed with paranoid personality disorder, we can average the values for each SWAP items to obtain a composite description of the prototypical paranoid patient.

A fortunate statistical consequence of averaging is that only SWAP items ranked highly for virtually all patients will have a high ranking in the composite description. If a descriptor does not apply to all or most patients in the sample, the item will not achieve a high score. Thus, examination of the highest-ranking items in the composite description for paranoid personality disorder reveals the *core psychological features* shared by paranoid patients treated in the community. This represents a purely empirical procedure for identifying the core features of a personality syndrome.

Method

A national sample of 530 experienced psychiatrists and clinical psychologists recruited from the rosters of the American Psychiatric Association and the American Psychological Association used the SWAP-200 to describe a current patient with a specified personality disorder diagnosis (for a more complete description of the study methods, see Shedler & Westen, 2004a). We aggregated the SWAP descriptions across all patients with a given personality disorder diagnosis to create a *composite description* for each personality disorder included in DSM-IV. The composite descriptions were highly reliable (coefficient alpha > 0.90 for all

¹⁰ The material in this section is adapted from Shedler and Westen, 2004a.

descriptions), indicating that the sample sizes were adequate to obtain stable and reproducible personality descriptions.¹¹

Results

We will describe the findings for a few personality disorders. (For a more complete account of the study findings for all ten personality disorders, see Shedler & Westen, 2004a.)

Cluster A: The “Odd” Cluster

Tables 1a-1c list the SWAP-200 items that received the highest scores or rankings in each composite description, along with the item’s mean score or ranking in the composite (indicating its centrality or importance in defining the personality disorder). Two findings are noteworthy. First, the descriptions differ systematically from the DSM-IV descriptions and include psychological features absent from the DSM criterion sets, especially items addressing inner life or intrapsychic experience. Second, there is considerable overlap in item content between the disorders. Thus, there are psychological features that are central to two or all three of the Cluster A disorders (e.g., difficulty making sense of other people’s behavior, problematic reality testing, a propensity to feel misunderstood or mistreated, a tendency toward social isolation). If we consider each composite description as a *whole* (that is, if we consider the “gist” or gestalt of the 15 to 20 most descriptive statements), the descriptions are readily distinguishable. However, if we limit the descriptions to just the first 8 to 9 items—the number included in DSM-IV criterion sets—it is more difficult to distinguish them. This suggests that criterion sets of 8 to 9 items are too small to provide PD descriptions that are both clinically accurate and adequately distinct (Shedler & Westen, 2004b; Westen & Shedler, 2000).

Paranoid Personality Disorder

Empirically observable features of paranoid personality disorder include aggression (“Tends to be angry or hostile”) and the defenses of externalization (“tends to blame own one’s failures or shortcomings on others”) and projection (“tends to see own unacceptable feelings or impulses in other people instead of in him/herself”). The findings are consistent with the psychodynamic hypothesis that projection of aggression is a central dynamic in paranoid personality (i.e.,

¹¹ The reliability of a composite or aggregate personality description is measured by coefficient alpha, which reflects the intercorrelations between the patients (columns of data) included in the aggregate description. The logic is identical to computing the reliability of a psychometric scale, except that patients are treated as scale “items” (columns in the data file) and SWAP-200 items are treated as cases (rows in the data file). See footnote 4 for additional details.

paranoid patients perceive the world as dangerous because they see their own hostility wherever they look). Similar findings emerged when we stratified the data by clinician theoretical orientation and omitted data provided by clinicians who described their theoretical orientation as psychoanalytic or psychodynamic (it is therefore highly unlikely that the reporting clinicians were simply describing their personality theories, rather than the observed characteristics of their patients). Other empirically observable characteristics of paranoid personality disorder absent from DSM-IV include feelings of victimization, difficulties understanding others' actions, hypersensitivity to slights, lack of close friendships and relationships, and the tendency for reasoning to become severely impaired under stress.

Cluster B: The "Dramatic" Cluster

Tables 2a-2d list the SWAP-200 items that received the highest ranking in the composite descriptions for the Cluster B disorders. Again, the descriptions of personality disorders differ systematically from the DSM-IV descriptions and place greater emphasis on inner life. Once again there is significant item overlap, but the disorders are readily distinguishable when the descriptions are considered in total.

Antisocial Personality Disorder

The composite description of antisocial patients includes multiple traits associated with the construct of psychopathy that preceded the current antisocial personality disorder diagnosis (Cleckley, 1941; Patrick & Zempolich, 1998). Included in the composite description, but absent from the DSM-IV criterion set, are items addressing lack of empathy, sadism, emotional manipulateness, imperviousness to consequences, and externalization of blame. In contrast, DSM-IV emphasizes behaviors associated with criminality (and would therefore miss the more "successful" psychopathic personalities who express their pathology in the world of business or politics).

Cluster C: The "Anxious" Cluster

Avoidant and Dependent Personality Disorder

The empirical portraits of avoidant and dependent personality disorders in Tables 3a and 3b help explain the excessive comorbidity between the disorders observed in virtually every study to date (Millon & Martinez, 1995; Westen & Shedler, 1999a). Patients diagnosed with these disorders share a *depressive* or *dysphoric core* that appears to pervade all areas of functioning. This depression or dysphoria is not captured by the current DSM criteria. Patients diagnosed by

their clinicians with avoidant personality disorder attempt to deal with dysphoria by keeping their distance from people, whereas those diagnosed with dependent personality disorder attempt to cope by clinging. However, both groups experience depression and despondency, feelings of inferiority, guilt, shame, anxiety, self-criticism, self-blame, passivity, and inhibitions. Clinicians appear to be using these diagnostic categories to describe patients who might be better conceptualized in terms of depressive personality disorder (see *Personality Patterns and Disorders*, p. 15).

Table 1a: Composite Description of Patients Diagnosed with Paranoid Personality Disorder

Item	Mean
Tends to feel misunderstood, mistreated, or victimized.	6.19
Is quick to assume that others want to harm or take advantage of him/her; tends to perceive malevolent intentions in others' words and actions.	5.97
Tends to be angry or hostile (whether consciously or unconsciously).	5.74
Tends to hold grudges; may dwell on insults or slights for long periods.	5.55
Tends to blame others for own failures or shortcomings; tends to believe his/her problems are caused by external factors.	5.26
Tends to avoid confiding in others for fear of betrayal; expects things she/he says or does will be used against him/her.	5.03
Tends to be critical of others.	5.03
Tends to react to criticism with feelings of rage or humiliation.	4.94
Lacks close friendships and relationships.	4.52
Tends to get into power struggles.	4.48
Has difficulty making sense of other people's behavior; often misunderstands, misinterprets, or is confused by others' actions and reactions.	4.48
Perception of reality can become grossly impaired under stress (e.g., may become delusional).	4.32
Tends to feel like an outcast or outsider; feels as if she/he does not truly belong.	4.26
Tends to express intense and inappropriate anger, out of proportion to the situation at hand.	4.23
Tends to feel helpless, powerless, or at the mercy of forces outside his/her control.	4.16
Tends to see own unacceptable feelings or impulses in other people instead of in him/herself.	4.03

Table 1b: Composite Description of Patients Diagnosed with Schizoid Personality Disorder

Item	Mean
Lacks close friendships and relationships.	5.85
Lacks social skills; tends to be socially awkward or inappropriate.	5.59
Appears to have a limited or constricted range of emotions.	5.44
Tends to feel like an outcast or outsider; feels as if she/he does not truly belong.	5.13
Tends to be inhibited or constricted; has difficulty allowing self to acknowledge or express wishes and impulses.	5.08
Tends to be shy or reserved in social situations.	4.95
Appearance or manner seems odd or peculiar (e.g., grooming, hygiene, posture, eye contact, speech rhythms, etc. seem somehow strange or "off").	4.56
Tends to avoid social situations because of fear of embarrassment or humiliation.	4.46
Has difficulty making sense of other people's behavior; often misunderstands, misinterprets, or is confused by others' actions and reactions.	4.31
Has difficulty acknowledging or expressing anger.	4.28
Tends to feel unhappy, depressed, or despondent.	4.23
Has difficulty allowing self to experience strong pleasurable emotions (e.g., excitement, joy, pride).	4.18
Tends to be passive and unassertive.	4.13
Appears to find little or no pleasure, satisfaction, or enjoyment in life's activities.	4.00
Tends to feel she/he is inadequate, inferior, or a failure.	3.97
Appears to have little need for human company or contact; is genuinely indifferent to the presence of others.	3.92
Appears inhibited about pursuing goals or successes; aspirations or achievements tend to be below his/her potential.	3.90
Tends to be anxious.	3.59

Table 1c: Composite Description of Patients Diagnosed with Schizotypal Personality Disorder

Item	Mean
Lacks close friendships and relationships.	6.17
Appearance or manner seems odd or peculiar (e.g., grooming, hygiene, posture, eye contact, speech rhythms, etc. seem somehow strange or “off”).	6.08
Reasoning processes or perceptual experiences seem odd and idiosyncratic (e.g., may make seemingly arbitrary inferences; may see hidden messages or special meanings in ordinary events).	5.17
Tends to feel like an outcast or outsider; feels as if she/he does not truly belong.	4.79
Lacks social skills; tends to be socially awkward or inappropriate.	4.79
Has difficulty making sense of other people’s behavior; often misunderstands, misinterprets, or is confused by others’ actions and reactions.	4.71
Perception of reality can become grossly impaired under stress (e.g., may become delusional).	4.63
Appears to have a limited or constricted range of emotions.	4.50
Tends to become irrational when strong emotions are stirred up; may show a noticeable decline from customary level of functioning.	4.08
Tends to be shy or reserved in social situations.	4.04
Tends to be anxious.	3.88
Tends to feel unhappy, depressed, or despondent.	3.83
Tends to feel helpless, powerless, or at the mercy of forces outside his/her control.	3.71
Tends to feel misunderstood, mistreated, or victimized.	3.58
Tends to avoid social situations because of fear of embarrassment or humiliation.	3.54
Has little psychological insight into own motives, behavior, etc.; is unable to consider alternate interpretations of his/her experiences.	3.54
Lacks a stable image of who she/he is or would like to become (e.g., attitudes, values, goals, and feelings about self may be unstable and changing).	3.50

Table 2a: Composite Description of Patients Diagnosed with Antisocial Personality Disorder

Item	Mean
Takes advantage of others; is out for number one; has minimal investment in moral values.	5.64
Tends to be deceitful; tends to lie or mislead.	5.50
Tends to engage in unlawful or criminal behavior.	5.36
Tends to be angry or hostile (whether consciously or unconsciously).	5.29
Has little empathy; seems unable to understand or respond to others' needs and feelings unless they coincide with his/her own.	5.04
Appears to experience no remorse for harm or injury caused to others.	4.93
Tends to blame others for own failures or shortcomings; tends to believe his/her problems are caused by external factors.	4.89
Tends to act impulsively, without regard for consequences.	4.89
Tends to show reckless disregard for the rights, property, or safety of others.	4.86
Tries to manipulate others' emotions to get what she/he wants.	4.75
Tends to be unconcerned with the consequences of his/her actions; appears to feel immune or invulnerable.	4.39
Tends to be unreliable and irresponsible (e.g., may fail to meet work obligations or honor financial commitments).	4.32
Has little psychological insight into own motives, behavior, etc.; is unable to consider alternate interpretations of his/her experiences.	4.21
Tends to get into power struggles.	4.07
Appears to gain pleasure or satisfaction by being sadistic or aggressive toward others	4.04
Tends to abuse alcohol.	4.04
Tends to be critical of others.	4.00
Tends to be conflicted about authority (e.g., may feel she/he must submit, rebel against, win over, defeat, etc.).	4.00
Tends to seek power or influence over others (whether in beneficial or destructive ways).	3.93
Has an exaggerated sense of self-importance.	3.75

Table 2b: Composite Description of Patients Diagnosed with Borderline Personality Disorder

Item	Mean
Emotions tend to spiral out of control, leading to extremes of anxiety, sadness, rage, excitement, etc.	5.05
Tends to feel unhappy, depressed, or despondent.	4.88
Tends to feel she/he is inadequate, inferior, or a failure.	4.42
Tends to fear she/he will be rejected or abandoned by those who are emotionally significant.	4.40
Is unable to soothe or comfort self when distressed; requires involvement of another person to help regulate affect.	4.28
Tends to feel helpless, powerless, or at the mercy of forces outside his/her control.	4.19
Tends to be angry or hostile (whether consciously or unconsciously).	4.05
Tends to be anxious.	4.05
Tends to react to criticism with feelings of rage or humiliation.	3.95
Tends to be overly needy or dependent; requires excessive reassurance or approval.	3.93
Tends to feel misunderstood, mistreated, or victimized.	3.79
Tends to become irrational when strong emotions are stirred up; may show a noticeable decline from customary level of functioning.	3.74
Tends to get into power struggles.	3.56
Tends to “catastrophize”; is prone to see problems as disastrous, unsolvable, etc.	3.51
Emotions tend to change rapidly and unpredictably.	3.51
Lacks a stable image of who she/he is or would like to become (e.g., attitudes, values, goals, and feelings about self may be unstable and changing).	3.49
Tends to feel like an outcast or outsider; feels as if she/he does not truly belong.	3.47
Tends to express intense and inappropriate anger, out of proportion to the situation at hand.	3.40

Table 2c: Composite Description of Patients Diagnosed with Histrionic Personality Disorder

Item	Mean
Expresses emotion in exaggerated and theatrical ways.	5.00
Tends to fear she/he will be rejected or abandoned by those who are emotionally significant.	4.66
Tends to be anxious.	4.43
Emotions tend to spiral out of control, leading to extremes of anxiety, sadness, rage, excitement, etc.	4.40
Tends to be overly needy or dependent; requires excessive reassurance or approval.	4.34
Tends to develop somatic symptoms in response to stress or conflict (e.g., headache, backache, abdominal pain, asthma, etc.).	3.77
Tends to get into power struggles.	3.63
Tends to become attached quickly or intensely; develops feelings, expectations, etc. that are not warranted by the history or context of the relationship.	3.60
Tends to be overly sexually seductive or provocative, whether consciously or unconsciously (may be inappropriately flirtatious, preoccupied with sexual conquest, prone to "lead people on," etc.).	3.60
Seeks to be the center of attention.	3.57
Tends to feel misunderstood, mistreated, or victimized.	3.54
Is articulate; can express self well in words.	3.46
Tends to become irrational when strong emotions are stirred up; may show a noticeable decline from customary level of functioning.	3.46
Tends to feel helpless, powerless, or at the mercy of forces outside his/her control.	3.37
Is unable to soothe or comfort self when distressed; requires involvement of another person to help regulate affect.	3.34
Emotions tend to change rapidly and unpredictably.	3.34
Tends to "catastrophize;" is prone to see problems as disastrous, unsolvable, etc.	3.29
Tends to feel unhappy, depressed, or despondent.	3.29
Tends to use his/her physical attractiveness to an excessive degree to gain attention or notice.	3.26
Tends to be angry or hostile (whether consciously or unconsciously).	3.17

Table 2d: Composite Description of Patients Diagnosed with Narcissistic Personality Disorder

Item	Mean
Appears to feel privileged and entitled; expects preferential treatment.	4.95
Has an exaggerated sense of self-importance.	4.68
Tends to be controlling.	4.53
Tends to be critical of others.	4.40
Tends to get into power struggles.	4.28
Tends to feel misunderstood, mistreated, or victimized.	4.28
Tends to be competitive with others (whether consciously or unconsciously).	4.25
Is articulate; can express self well in words.	4.25
Tends to react to criticism with feelings of rage or humiliation.	4.22
Tends to be angry or hostile (whether consciously or unconsciously).	4.15
Has little empathy; seems unable to understand or respond to others' needs and feelings unless they coincide with his/her own.	4.10
Tends to blame others for own failures or shortcomings; tends to believe his/her problems are caused by external factors.	4.00
Seeks to be the center of attention.	3.63
Tends to be arrogant, haughty, or dismissive.	3.63
Seems to treat others primarily as an audience to witness own importance, brilliance, beauty, etc.	3.50
Has fantasies of unlimited success, power, beauty, talent, brilliance, etc.	3.43
Tends to hold grudges; may dwell on insults or slights for long periods.	3.40
Expects self to be "perfect" (e.g., in appearance, achievements, performance, etc.).	3.38

Table 3a: Composite Description of Patients Diagnosed with Avoidant Personality Disorder

Item	Mean
Tends to feel she/he is inadequate, inferior, or a failure.	6.34
Tends to be shy or reserved in social situations.	6.26
Tends to avoid social situations because of fear of embarrassment or humiliation.	5.94
Tends to feel ashamed or embarrassed.	5.71
Tends to be anxious.	5.60
Tends to feel like an outcast or outsider; feels as if she/he does not truly belong.	5.51
Tends to be inhibited or constricted; has difficulty allowing self to acknowledge or express wishes and impulses.	5.31
Tends to be passive and unassertive.	5.29
Tends to feel unhappy, depressed, or despondent.	5.20
Tends to be self-critical; sets unrealistically high standards for self and is intolerant of own human defects.	4.91
Lacks close friendships and relationships.	4.89
Tends to blame self or feel responsible for bad things that happen.	4.86
Tends to fear she/he will be rejected or abandoned by those who are emotionally significant.	4.83
Tends to feel guilty.	4.77
Lacks social skills; tends to be socially awkward or inappropriate.	4.74
Appears inhibited about pursuing goals or successes; aspirations or achievements tend to be below his/her potential.	4.49

Table 3b: Composite Description of Patients Diagnosed with Dependent Personality Disorder

Item	Mean
Tends to be overly needy or dependent; requires excessive reassurance or approval.	6.13
Tends to fear she/he will be rejected or abandoned by those who are emotionally significant.	5.55
Tends to feel she/he is inadequate, inferior, or a failure.	5.47
Tends to feel unhappy, depressed, or despondent.	5.26
Tends to be ingratiating or submissive (e.g., may consent to things she/he does not agree with or does not want to do, in the hope of getting support or approval).	5.24
Tends to feel helpless, powerless, or at the mercy of forces outside his/her control.	5.16
Tends to feel guilty.	4.89
Tends to be passive and unassertive.	4.76
Tends to be anxious.	4.55
Tends to blame self or feel responsible for bad things that happen.	4.53
Has difficulty acknowledging or expressing anger.	4.53
Tends to feel ashamed or embarrassed.	4.39
Is unable to soothe or comfort self when distressed; requires involvement of another person to help regulate affect.	4.37
Has trouble making decisions; tends to be indecisive or to vacillate when faced with choices.	4.26
Appears inhibited about pursuing goals or successes; aspirations or achievements tend to be below his/her potential.	4.21
Tends to express aggression in passive and indirect ways (e.g., may make mistakes, procrastinate, forget, become sulky, etc.).	4.03
Tends to get drawn into or remain in relationships in which she/he is emotionally or physically abused.	3.79

Table 3c: Composite Description of Patients Diagnosed with Obsessive-Compulsive Personality Disorder

Item	Mean
Tends to be conscientious and responsible.	5.83
Tends to be self-critical; sets unrealistically high standards for self and is intolerant of own human defects.	5.20
Has moral and ethical standards and strives to live up to them.	5.17
Tends to be overly concerned with rules, procedures, order, organization, schedules, etc.	4.89
Tends to be anxious.	4.86
Tends to be controlling.	4.80
Tends to become absorbed in details, often to the point that she/he misses what is significant in the situation.	4.74
Expects self to be "perfect" (e.g., in appearance, achievements, performance, etc.).	4.69
Tends to blame self or feel responsible for bad things that happen.	4.49
Tends to feel guilty.	4.43
Tends to adhere rigidly to daily routines and become anxious or uncomfortable when they are altered.	4.29
Is troubled by recurrent obsessional thoughts that she/he experiences as senseless and intrusive.	4.26
Is articulate; can express self well in words.	4.26
Tends to be inhibited or constricted; has difficulty allowing self to acknowledge or express wishes and impulses.	4.14
Is excessively devoted to work and productivity, to the detriment of leisure and relationships.	4.11
Tends to feel unhappy, depressed, or despondent.	4.09
Has difficulty allowing self to experience strong pleasurable emotions (e.g., excitement, joy, pride).	3.97

Discussion of Empirical Findings

Advantages of Expanded Criterion Sets

A consistent theme running through the findings is that DSM-IV criterion sets are too narrow. They do not capture the richness and complexity of the personality syndromes observed empirically in patients treated in the community, nor do they capture the complexity of personality disorders as they are defined by DSM-IV itself. The preamble to Axis II defines personality disorders in terms of multiple domains of functioning including cognition, affectivity, interpersonal relations, and impulse regulation. However, the PD criterion sets do not actually encompass these domains of functioning (Millon, 1990; Millon & Davis, 1997).

DSM-IV limits the number of diagnostic criteria to 8 or 9 criteria (items) per disorder, but it is clinically and psychometrically impossible for such small item sets simultaneously to describe personality syndromes in their complexity, and to describe distinct (non-overlapping) syndromes. Certain traits play central roles in more than one personality disorder (e.g., lack of empathy is characteristic of both narcissistic and antisocial personality disorder; hostility is characteristic of paranoid, antisocial, borderline, and narcissistic personality disorders). Excluding these traits from the PD criterion sets leads to clinically inaccurate descriptions, but including the same items in multiple criterion sets leads to excess comorbidity (i.e., low specificity). As now constituted, Axis II cannot transcend this paradox.

The paradox could be resolved by (1) expanding the size of the criterion sets, and (2) diagnosing personality disorders as configurations or gestalts rather than by tabulating individual symptoms (for discussion of such an approach to diagnosis, which we call “prototype matching,” see Shedler & Westen, 2004a; Westen & Shedler, 2000). For example, the composite descriptions of narcissistic and antisocial personality disorder contain numerous overlapping traits, yet they are conceptually distinct and would be difficult to confuse. Expanding the size of the criterion sets would (1) help bridge the gap between science and practice by making the descriptions of personality disorders in the DSM more faithful to clinical reality, (2) make the personality disorder descriptions more faithful to the conceptual definition of *personality disorder* (i.e., multifaceted syndromes encompassing multiple domains of functioning), and (3) reduce comorbidities among personality disorder diagnoses by making the diagnostic categories more distinct from one another.

Addressing Intrapsychic Processes and Inner Experience

DSM-IV tends to underemphasize inner experience or intrapsychic processes that are centrally defining of personality disorders, which limits both its clinical relevance and its empirical fidelity. For example, the data strongly indicate that

aggression and the defenses of externalization and projection are defining features of paranoid personality disorder, yet they are not included in the DSM-IV criterion set. The data indicate that hostility, sadism, lack of empathy, lack of insight, self-importance, and power-seeking are defining of antisocial personality disorder. However, these aspects of mental life are absent from the DSM description, which instead emphasizes behavioral markers such as criminality and lack of stable employment. Feelings of inadequacy and inferiority, shame, embarrassment, passivity, depression, anxiety, self-blame, and guilt appear centrally defining of both avoidant and dependent personality disorder (which should probably be subsumed by a diagnosis of depressive personality disorder); instead, DSM-IV emphasizes behavioral indicators of social avoidance in the former and dependency in the latter.

Identifying Optimal Diagnostic Groupings

This chapter has focused on the diagnostic categories currently included in DSM-IV, but the findings raise broader questions about whether these categories are the optimal ones. For example, the composite descriptions of avoidant and dependent personality disorder overlap substantially and contain numerous features that would be better characterized in terms of a depressive personality syndrome (e.g., the tendency to feel unhappy, depressed, despondent; to feel inadequate, inferior, or a failure; to blame themselves for bad things that happen; to be inhibited about pursuing goals or successes; to feel ashamed or embarrassed; to fear rejection and abandonment; etc.). A depressive personality disorder category deserves consideration for DSM-V.

CONCLUSION: INTEGRATING SCIENCE AND PRACTICE

A clinically useful diagnostic system should encompass the spectrum of personality pathology seen in clinical practice and have meaningful implications for treatment. An empirically sound diagnostic system should facilitate reliable and valid diagnoses: Independent clinicians should be able to arrive at the same diagnosis, the diagnoses should be relatively distinct from one another, and each diagnosis should be associated with unique and theoretically meaningful correlates, antecedents, and sequelae (Livesley & Jackson, 1992; Millon, 1991; Robins & Guze, 1970).

One obstacle to achieving this ideal has been an unfortunate schism in the mental health professions between science and practice. Too often, research has been conducted in isolation from the crucial data of clinical observation. The results often strike clinicians as naïve and of dubious clinical relevance. Ultimately, the most empirically elegant diagnostic system will have little impact if clinicians do not find it helpful for understanding their patients (Shedler &

Westen, 2005). On the other hand, clinical theory has too often developed with little regard for questions of falsifiability or empirical credibility. The results have often struck researchers as scientifically naïve.

The SWAP represents an effort to bridge the schism between science and practice by quantifying clinical wisdom and expertise and making clinical constructs accessible to empirical study. It relies on clinicians to do what they do best, namely, making observations and inferences about individual patients they know and treat. It relies on quantitative methods to do what they do best, namely aggregating observations to discern relationships and commonalities, and combining data to yield optimal predictions (cf. Sawyer, 1966). The findings raise possibilities for developing a classification of personality disorders that is both empirically sound and clinically (and psychodynamically) meaningful; for integrating descriptive psychiatric diagnosis with clinical case formulation; for assessing personality change (not just symptom remission) in psychotherapy; and for assessing individual patients in ways that integrate the best features of clinical judgment and psychometric rigor. The SWAP attempts to provide a “language” for case description that is at once clinically rich enough to describe the complexities of the patients we treat, and empirically rigorous enough to meet the requirements of science. There remains a sizeable schism between clinical practitioners and empirical researchers. Perhaps this new language will be a step toward one that all parties can speak.

REFERENCES

- Blais, M., & Norman, D. (1997). A psychometric evaluation of the DSM-IV Personality Disorder Criteria. *Journal of Personality Disorders*, 11, 168-176.
- Blashfield, R. (1985). Exemplar prototypes of personality disorder diagnoses. *Comprehensive Psychiatry*, 26, 11-21.
- Block, J. (1971). *Lives through time*. Berkeley, CA: Bancroft.
- Block, J. (1978). *The Q-sort method in personality assessment and psychiatric research*. Palo Alto, CA: Consulting Psychologists Press.
- Bradley, R., & Westen, D. (2003). *Validity of SWAP-200 personality diagnosis in an outpatient sample*. Unpublished manuscript. Emory University.
- Cantor, N., & Genero, N. (1986). Psychiatric diagnosis and natural categorization: A close analogy. In Contemporary directions in psychopathology. In G.L. Klerman (Ed.), *Toward the DSM-IV* (pp. 233-256). New York: Guilford.
- Clark, L. (1992). Resolving taxonomic issues in personality disorders: The value of larger scale analyses of symptom data. *Journal of Personality Disorders*, 6, 360-376.
- Cleckley, H. (1941). *The mask of sanity*. St. Louis, MO: Mosby Co.
- Davis, R., Blashfield, R., & McElroy, R. (1993). Weighting criteria in the diagnosis of a personality disorder: A demonstration. *Journal of Abnormal Psychology*, 102, 319-322.
- First, M., Spitzer, R., Gibbon, M., Williams, J., Davies, J. B., Howes, M., et al. (1995). The Structured Clinical Interview for DSM-III-R Personality Disorders (SCID-II). Part II: Multi-site test-retest reliability study. *Journal of Personality Disorders*, 9, 92-104.

- Grilo, C. M., Sanislow, C.A., & McGlashan, T. H. (2002). Co-occurrence of DSM-IV personality disorders with borderline personality disorder. *Journal of Nervous and Mental Disease*, 190, 552-553.
- Gunderson, J. G. (2001). *Borderline personality disorder: A clinical guide*. Washington, DC: American Psychiatric Publishing, Inc.
- 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.
- Harkness, A. (1992). Fundamental topics in the personality disorders: candidate trait dimensions from lower regions of the hierarchy. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, 4, 251-259.
- John, O. (1990). The "big five" factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In L. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 66-100). New York: Guilford Press.
- Kernberg, O. (1975). *Borderline conditions and pathological narcissism*. New York: Jason Aronson.
- Kernberg, O. (1984). *Severe personality disorders*. New Haven, CT: Yale University Press.
- Kim, N. S., & Ahn, W. (2002). Clinical psychologists' theory-based representations of mental disorders predict their diagnostic reasoning and memory. *Journal of Experimental Psychology*, 131(4), 451-476.
- Klonsky, E. D., Oltmanns, T. F., & Turkheimer, E. (2002). Informant-reports of personality disorder: Relation to self-reports and future research directions. *Clinical Psychology: Science and Practice*, 9, 300-311.
- Kohut, H. (1971). *The analysis of the self*. New York: International Universities Press.
- Linehan, M. M. (1993). *Cognitive-behavioral treatment of borderline personality disorder*. New York: Guilford.
- Lingiardi V., Shedler, J., & Gazillo, F. (2006). Assessing personality change in psychotherapy with the SWAP-200: A case study. *Journal of Personality Assessment*, 86, 23-32.
- Livesley, W. J. (Ed.). (1995). *The DSM-IV personality disorders*. New York: Guilford Press.
- Livesley, W. J., & Jackson, D. N. (1992) Guidelines for developing, evaluating, and revising the classification of personality disorders. *Journal of Nervous and Mental Disease*, 180, 609-618.
- Main, M., Kaplan, N., & Cassidy, J. (1985). Security in infancy, childhood, and adulthood: A move to the level of representation. *Monographs of the Society for Research in Child Development*, 50(1-2, Serial No. 209).
- Marin-Avellan, L., McGauley, G., Campbell, C., & Fonagy, P. (2005). Using the SWAP-200 in a personality-disordered forensic population: Is it valid, reliable and useful? *Journal of Criminal Behaviour and Mental Health*, 15, 28-45.
- McCrae, R., & Costa, P. (1990). *Personality in adulthood*. New York: Guilford Press.
- McWilliams, N. (1999). *Psychoanalytic case formulation*. New York: Guilford Press.
- Millon, T. (1990). *Toward a new psychology*. New York: Wiley.
- Millon, T. (1991). Classification in psychopathology: Rationale, alternatives and standards. *Journal of Abnormal Psychology*, 100, 245-261.
- Millon, T., & Davis, R. D. (1997). The place of assessment in clinical science. In T. Millon (Ed.), *The Millon Inventories: Clinical and personality assessment* (pp. 3-20). New York: Guilford Press.
- Millon, T. & Martinez, A. (1995). Avoidant personality disorder. In W. J. Livesley (Ed.), *The DSM-IV personality disorders* (pp. 218-233). New York: Guilford Press.
- Morey, L. C. (1988). Personality disorders in DSM-III and DSM-III-R: convergence, coverage, and internal consistency. *American Journal of Psychiatry*, 145, 573-577.
- Morey, L. C. (1991). *The Personality Assessment Inventory: Professional manual*. Odessa, FL: Psychological Assessment Resources.

- Oldham, J., Skodol, A., Kellman, H. D., Hyler, S., Rosnick, L., & Davies, M. (1992). Diagnosis of DSM-III-R personality disorders by two semistructured interviews: Patterns of comorbidity. *American Journal of Psychiatry*, 149, 213-220.
- Patrick, C. J., & Zempolich, K. A. (1998). Emotion and aggression in the psychopathic personality. *Aggression and violent behavior*, 3(4), 303-338.
- Perry, J. C. (1992). Problems and considerations in the valid assessment of personality disorders. *American Journal of Psychiatry*, 149, 1645-1653.
- Perry, J. C., & Cooper, S. H. (1987). Empirical studies of psychological defense mechanisms. In J. Cavenar & R. Michels (Eds.), *Psychiatry*. Philadelphia: JB Lippincott.
- Pilkonis, P. A., Heape, C. L., Proietti, J. M., Clark, S. W., McDavid, J. D., & Pitts, T. E. (1995). The reliability and validity of two structured diagnostic interviews for personality disorders. *Archives of General Psychiatry*, 52, 1025-1033.
- Pilkonis, P. A., Heape, C. L., Ruddy, J., & Serrao, P. (1991). Validity in the diagnosis of personality disorders: The use of the LEAD standard. *Psychological Assessment*, 31, 46-54.
- Robins, E., & Guze, S. (1970). The establishment of diagnostic validity in psychiatric illness: Its application to schizophrenia. *American Journal of Psychiatry*, 126, 983-987.
- Sawyer, J. (1966). Measurement and prediction, clinical and statistical. *Psychological Bulletin*, 66, 178-200.
- Shedler, J. (2002). A new language for psychoanalytic diagnosis. *Journal of the American Psychoanalytic Association*, 50(2), 429-456.
- Shedler, J., Mayman, M., & Manis, M. (1993). The illusion of mental health. *American Psychologist*, 48, 1117-1131.
- Shedler, J., & Westen, D. (1998). Refining the measurement of Axis II: A Q-sort procedure for assessing personality pathology. *Assessment*, 5(4), 333-353.
- Shedler, J., & Westen, D. (2004a). Refining DSM-IV personality disorder diagnosis: Integrating science and practice. *American Journal of Psychiatry*, 161, 1350-1365.
- Shedler, J., & Westen, D. (2004b). Dimensions of personality pathology: An alternative to the Five Factor Model. *American Journal of Psychiatry*, 161, 1743-1754.
- Shedler, J., & Westen, D. (2005). A simplistic view of the Five Factor Model: Drs. Shedler and Westen reply. *American Journal of Psychiatry*, 162, 1551.
- Skodol, A., Oldham, J., Rosnick, L., Kellman, D., & Hyler, S. (1991). Diagnosis of DSM-III-R personality disorders: A comparison of two structured interviews. *International Journal of Methods in Psychiatric Research*, 1, 13-26.
- Spitzer, R. L. (1983). Psychiatric diagnosis: Are clinicians still necessary? *Comprehensive Psychiatry*, 24, 399-411.
- Vaillant, G. (Ed.) (1992). *Ego mechanisms of defense: A guide for clinicians and researchers*. Washington, DC: American Psychiatric Press.
- Walker, E., & Lewine, R. J. (1990). Prediction of adult-onset schizophrenia from childhood home movies of the patients. *American Journal of Psychiatry*, 147(8), 1052-1056.
- Watson, D., & Sinha, B. K. (1998). Comorbidity of DSM-IV personality disorders in a nonclinical sample. *Journal of Clinical Psychology*, 54(6), 773-780.
- Westen, D. (1991). Social cognition and object relations. *Psychology Bulletin*, 109, 429-455.
- Westen, D. (1997). Divergences between clinical and research methods for assessing personality disorders: Implications for research and the evolution of Axis II. *American Journal of Psychiatry*, 154, 895-903.
- Westen, D. (2002). *Clinical Diagnostic Interview*. Unpublished manual. Emory University. Available from www.psychsystems.net/lab. (Accessed 1/30/06.)
- Westen, D., & Arkowitz-Westen, L. (1998). Limitations of Axis II in diagnosing personality pathology in clinical practice. *American Journal of Psychiatry*, 155, 1767-1771.

- Westen, D., Heim, A.K., Morrison, K., Patterson, M., & Campbell, L. (2002). Classifying and diagnosing psychopathology: A prototype matching approach. In M. Malik (Ed.), *Rethinking the DSM: Psychological perspectives* (pp. 221-250). Washington DC: American Psychological Association Press.
- Westen, D., Lohr, N., Silk, K., Gold, L., & Kerber, K. (1990). Object relations and social cognition in borderlines, major depressives, and normals: A TAT analysis. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, 2, 355-364.
- Westen, D., & Muderrisoglu, S. (2003). Reliability and validity of personality disorder assessment using a systematic clinical interview: Evaluating an alternative to structured interviews. *Journal of Personality Disorders*, 17, 350-368.
- Westen, D., & Muderrisoglu, S. (in press). Clinical assessment of pathological personality traits. *American Journal of Psychiatry*.
- Westen, D., Muderrisoglu, S., Fowler, C., Shedler, J., & Koren, D. (1997). Affect regulation and affective experience: Individual differences, group differences, and measurement using a Q-sort procedure. *Journal of Consulting and Clinical Psychology*, 65, 429-439.
- Westen, D., & Shedler, J. (1999a). Revising and assessing Axis II: I. Developing a clinically and empirically valid assessment method. *American Journal of Psychiatry*, 156(2), 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(2), 258-272.
- Westen, D., & Shedler, J. (2000). A prototype matching approach to diagnosing personality disorders toward DSM-V. *Journal of Personality Disorders*, 14(2), 109-126.
- 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.
- Widiger, T. A. (1993). The DSM-III-R categorical personality disorder diagnoses: A critique and an alternative. *Psychological Inquiry*, 4(2), 75-90.
- Widiger, T. A., & Corbitt, E. M. (1995). Antisocial personality disorder. In J. W. Livesley (Ed.), *The DSM-IV personality disorders. Diagnosis and treatment of mental disorders* (pp. 103-126). New York: Guilford Press.
- Widiger, T. A., & Simonsen, E. S. (1995). Alternative dimensional models of personality disorder: Finding a common ground. *Journal of Personality Disorders*, 19, 110-130.
- Wilkinson-Ryan, T., & Westen, D. (2000). Identity disturbance in borderline personality disorder: An empirical investigation. *American Journal of Psychiatry*, 157, 528-541.
- Zimmerman, M. (1994). Diagnosing personality disorders: A review of issues and research methods. *Archives of General Psychiatry*, 51, 225-245.

