

9

INTEGRATING CLINICAL AND EMPIRICAL PERSPECTIVES ON PERSONALITY: THE SHEDLER–WESTEN ASSESSMENT PROCEDURE (SWAP)

JONATHAN SHEDLER

It is well known that [Paul Meehl] not only thinks it important for a psychologist to work as a responsible professional with real-life clinical problems but, further, considers the purely “theoretical” personality research of academic psychologists to be unusually naïve and unrealistic when the researcher is not a seasoned, practicing clinician. (Meehl, 1973)

One of the greatest challenges facing psychology and psychiatry is the schism between science and practice. The schism is especially pronounced when it comes to understanding personality. For skilled clinical practitioners, personality assessment generally means *clinical case formulation*: understanding the patterns of thinking, feeling, fantasizing, desiring, fearing, coping, defending, attaching, relating, experiencing self and others, and so on, that make a person unique and (if he is in treatment) underlie his suffering. Understanding personality in this way requires skill. Sophisticated clinicians consider a range of information, attending not only to what patients say but also to how they say it, drawing inferences from patients’ accounts of their lives and relationships, from their manner of interacting with the clinician, and from their own emotional reactions

This chapter contains ideas first published in Shedler & Westen, 2006. Address e-mail correspondence to jonathan@shedler.com.

<http://dx.doi.org/10.1037/14549-010>

Personality Disorders: Toward Theoretical and Empirical Integration in Diagnosis and Assessment,
S. K. Huprich (Editor)

Copyright © 2015 by Jonathan Shedler.

to the patient (Betan, Heim, Conklin, & Westen, 2005; McWilliams, 2011; Peebles, 2012).

For example, clinicians generally do not assess lack of empathy, a central feature of narcissistic personality, by administering self-report questionnaires or asking direct questions. A moment's reflection reveals the futility of doing so: It would be a rare narcissistic patient who could report his own lack of empathy. More likely, the patient would describe himself as a caring person and a wonderful friend. An initial sign of lack of empathy on the part of the patient may be a subtle sense on the part of the clinician of being interchangeable or replaceable, or of being treated as a sounding board rather than as a person, or of feeling put down (Betan et al., 2005; Kernberg, 1975, 1984; McWilliams, 2011). Lack of empathy is something clinicians *infer*.

In other words, the clinician's subjective experience of the patient is a source of data and a vehicle for hypothesis generation. The clinician might go on to consider whether she often feels this way with this patient and whether such feelings are characteristic for her in her role as therapist. She might then become aware that the patient tends to describe other people more in terms of the functions they serve or needs they fulfill than in terms of who they are as people. She might further consider how these observations dovetail with the facts the patient has provided about his life and the problems that brought him to treatment. This kind of thinking lies at the heart of clinically sophisticated approaches to personality.

It is just such clinical judgment and inference that research approaches have eschewed. With respect to descriptive psychiatry, successive editions of the *Diagnostic and Statistical Manual of Mental Disorders (DSM)* have minimized the role of clinical inference and treated personality diagnosis as a largely technical task of tabulating signs and symptoms (for a discussion of the limitation of *DSM* personality disorder diagnosis, see Shedler & Westen, 2007).

With respect to academic psychology, personality research has focused on dimensional trait models derived from factor analysis of questionnaire data, notably the five-factor model (FFM) and its variants (e.g., Widiger & Simonsen, 2005). Although they derive from different traditions, academic psychology and descriptive psychiatry have both sought to minimize if not eliminate clinical inference and deduction. Indeed, the FFM was developed without input from clinical practitioners or theorists. It derives instead from questionnaire responses of laypersons describing themselves and their social acquaintances. Although the model is valuable for certain purposes, many experienced clinicians see it as quite removed from their clinical needs (Kernberg, 1996; Rottman, Ahn, Sanislow, & Kim, 2009; Shedler et al., 2010; Spitzer,

First, Shedler, Westen, & Skodal, 2008). Proposals by academic psychologists to replace clinical personality concepts with the FFM betray a devaluing attitude toward clinical knowledge and exemplify the science–practice schism. This is what Paul Meehl decried in the epigraph that begins this chapter.

ON THE SCIENCE–PRACTICE SCHISM

There is no inherent reason why the mental health professions must choose between clinical depth and scientific rigor. Good clinical work depends on scientific thinking and reasoning; it is characterized by an ongoing, cyclical process of data collection, hypothesis generation, hypothesis testing, and hypothesis revision. Empirical research rests (one hopes) on psychologically sophisticated human judgment from beginning to end, starting with judgments about what psychological phenomena are important to study, through judgments about how to conceptualize and operationalize them, to judgments about how to revise hypotheses as new information emerges.

Thus, good science and good clinical work involve a reciprocal interplay between the observations and judgments that lead to sound hypotheses and the investigation necessary to test them (in the language of philosopher of science Hans Reichenbach, 1938, a *context of discovery* and a *context of justification*). At its worst, clinical personality theory can seem a context of discovery without a credible context of justification—in other words, psychologically rich inference and conjecture without a credible basis for sifting sound from unsound ideas. At its worst, academic personality research can seem a context of justification without a credible context of discovery—in other words, sophisticated methodological tools applied to psychologically vacuous ideas.

Diagnosis and Case Formulation, Clinical and Statistical

The solution to the science–practice schism cannot be to turn back the clock and abandon scientific advances of the past decades, nor can it be to ignore more than a century of cumulative clinical observation and knowledge. Efforts to eliminate clinical observation and inference from empirical personality research may, inadvertently, exclude psychological phenomena of paramount importance (Cousineau & Shedler, 2006; Shedler, Mayman, & Manis, 1993).

The approach to personality described in this chapter, based on the Shedler–Westen Assessment Procedure (SWAP), bridges clinical and

empirical approaches to personality and seeks to integrate the strengths of each. The approach relies on clinicians to do what clinicians do well: observe and describe individual patients or clients they know. It relies on statistical methods to do what they do well: combine information in optimal ways to derive maximally valid information (Meehl, 1954; Sawyer, 1966; Westen & Weinberger, 2004). The goal is to provide a means of conceptualizing and assessing personality that is both clinically relevant and scientifically sound.

The remainder of this chapter (a) discusses the challenges of using clinical observation and inference in research, (b) describes the development of the SWAP as a method for systematizing clinical observation and understanding, (c) illustrates its use for diagnosis and clinical case formulation, and (d) describes a new taxonomy for personality diagnosis that is both empirically based and clinically relevant.

The Challenge of Clinical Data

It has become a truism that “clinical judgment is unreliable,” but truisms are not truths. In fact, the problem with clinical observation and inference is not that they are unreliable, as many investigators are happy to repeat (for a review, see Westen & Weinberger, 2004). The problem, rather, is that they come in a form that is difficult to work with. Rulers measure in inches and scales measure in pounds, but what metric do clinicians share? Consider three clinicians describing the same case. One might speak of schemas and beliefs, another of learning and conditioning, and the third, perhaps, of conflict and defense. It is not readily apparent whether the clinicians can or cannot make the same observations and inferences. There are three possibilities: (a) They may be observing the same thing but using different language and metaphor systems to describe it; (b) they may be attending to different aspects of the clinical material, as in the parable of the elephant and the blind men; or (c) they may not be able to make the same observations at all. To find out whether the clinicians can make the same observations and inferences, investigators must first ensure that they are speaking the same language and attending to the same spectrum of clinical phenomena.

DEVELOPING A STANDARD VOCABULARY FOR CASE DESCRIPTION

The SWAP is a tool for personality diagnosis and clinical case formulation that provides clinicians of all theoretical orientations with a standard vocabulary for clinical case description (Shedler & Westen, 2004a, 2004b, 2007; Westen & Shedler, 1999a, 1999b; Westen, Shedler, Bradley, & DeFife,

2012). The vocabulary consists of 200 personality-descriptive statements, each of which may describe a given patient very well, somewhat, or not at all. The clinician describes a patient by ranking the statements into eight categories, from most descriptive of the patient (scored 7) to not descriptive or irrelevant (scored 0). Thus, the SWAP yields a score from 0 to 7 for 200 personality-descriptive variables.

The major editions of the SWAP instrument are the SWAP-200 and the newer, revised SWAP-II (their precursor was the SWAP-167; Shedler & Westen, 1998). In this chapter, I use the acronym SWAP to refer to concepts and findings that apply to both major editions; I specify SWAP-200 or SWAP-II when a finding applies to a specific edition. The instrument can be obtained from <http://www.SWAPassessment.org>.¹

The “standard vocabulary” of the SWAP allows clinicians to provide comprehensive, in-depth psychological descriptions of patients in a form that is systematic and quantifiable, and it ensures that all clinicians attend to the same spectrum of clinical phenomena. SWAP statements stay close to the clinical data (e.g., “Tends to get into power struggles” or “Is capable of sustaining meaningful relationships characterized by genuine intimacy and caring”), and statements that require inference or deduction are written in clear, jargon-free language (e.g., “Tends to express anger in passive and indirect ways [e.g., may make mistakes, procrastinate, forget, become sulky, etc.]” or “Tends to see own unacceptable feelings or impulses in other people instead of in him/herself”).

SWAP Item Set

The initial SWAP item pool was drawn from a wide range of sources, including the clinical literature on personality pathology written over the past 50 years (e.g., Kernberg, 1975, 1984; Kohut, 1971; Linehan, 1993; Shapiro, 1965); DSM Axis II diagnostic criteria included in of the third through fourth editions of the DSM (American Psychiatric Association, 1980, 1994); selected DSM Axis I criteria that could reflect enduring dispositions (e.g., depression and anxiety); research on coping, defense, and affect regulation (e.g., Perry & Cooper, 1987; Shedler et al., 1993; Vaillant, 1992; Westen, Muderrisoglu, Fowler, Shedler, & Koren, 1997); research on interpersonal functioning in patients with PDs (e.g., Westen, 1991; Westen, Lohr, Silk, Gold, & Kerber, 1990); research on personality traits in nonclinical populations (e.g., Block,

¹Versions of the SWAP have been developed for adolescent personality assessment as well (e.g., Westen, Dutra, & Shedler, 2005; Westen, Shedler, Durrett, Glass, & Martens, 2003), but discussion of adolescent personality and its assessment is beyond the scope of this chapter.

1971; John, 1990; McCrae & Costa, 1990); research on personality pathology conducted since the development of *DSM* Axis II (see, e.g., Livesley, 1995); pilot studies in which observers watched videotaped interviews of patients with personality pathology and described them using draft versions of the SWAP item set; and the clinical experience of the SWAP authors.

Perhaps most important, the current SWAP item set is the product of a 12-year iterative item revision process that incorporated the feedback of more than 2,000 clinician-consultants of all theoretical orientations who used earlier versions of the instrument 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 (Westen et al., 2012).

The methods used to develop and refine the SWAP item set ensured the inclusion of clinically crucial concepts that are absent from other personality item sets. For example, clinical theorists have identified the phenomena of *splitting* and *projective identification* as central, defining features of borderline personality (Clarkin, Yeomans, & Kernberg, 2006; Gabbard, 2005; Kernberg, 1975, 1984; McWilliams, 2011). These concepts are notably absent both from the *DSM* and from dimensional trait models of personality. The SWAP-II addresses splitting with items like “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.)” and “Expresses contradictory feelings or beliefs without being disturbed by the inconsistency; has little need to reconcile or resolve contradictory ideas.” It addresses projective identification with items like “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)” and “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.)”

I provide these examples of SWAP items only to illustrate that it is possible to conduct empirical research without sacrificing clinical richness and complexity and to operationalize clinical (in this instance, psychodynamic) concepts and theories that many investigators dismiss as not lending themselves to empirical investigation. I am not (yet) addressing the question of whether the theories are correct. Rather, I am making the point that such concepts, which reflect the accrued experience of generations of skilled clinical observers, deserve to be taken seriously, as hypotheses to test. Neither

DSM-based structured interviews nor FFM instruments could ever confirm or disconfirm these hypotheses *because they make no attempt to address them*.

The methods used to develop and refine the SWAP item set were successful in creating a relatively comprehensive vocabulary for clinical case description. In a sample of 1,201 psychologists and psychiatrists who used the SWAP–II to describe a current patient, 84% “agreed” or “strongly agreed” that “the SWAP–II allowed me to express the things I consider important about my patient’s personality” (fewer than 5% disagreed).

Scoring the SWAP

The SWAP is based on the Q-sort method, which requires assessors to assign each score (0–7) a specified number of times (i.e., there is a “fixed” score distribution). The fixed score distribution is asymmetric, with many items receiving low scores and progressively fewer items receiving higher scores (the shape of the fixed distribution mirrors the naturally occurring distribution in the population; for a discussion of this and other psychometric issues, see Westen & Shedler, 2007). Use of a fixed distribution has psychometric advantages and reduces measurement error or “noise” inherent in standard rating scales.² The method maximizes the opportunity to observe statistical relations where they exist but does not, as some have mistakenly speculated, inflate or otherwise impact reliability or validity coefficients. Both Monte Carlo simulations and empirical data demonstrate that the shape of the SWAP fixed score distribution has no effect on the magnitude of correlation coefficients (Blagov, Bi, Shedler, & Westen, 2012). The psychometric rationale for the Q-sort method has been described in detail by Block (1961/1978).

When the SWAP is used in the context of psychotherapy, an experienced clinician can score the instrument after a minimum of 6 clinical contact hours with a patient. If a patient or subject is seen for assessment only—for example, in research, forensic, or personnel assessment contexts—the SWAP can be scored on the basis of the Clinical Diagnostic Interview (CDI), which systematizes and compresses into an approximately 2.5-hour time frame the kind of interviewing most skilled clinicians engage in during the initial hours of patient contact to assess personality (Westen, 2004; Westen & Muderrisoglu, 2003; Westen & Weinberger, 2004). The SWAP

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

can also be scored reliably from other, comparably psychologically rich interview sources (e.g., Marin-Avellan, McGauley, Campbell, & Fonagy, 2005).

Capturing Clinical Nuance

Just as academic researchers tend to be skeptical regarding clinical observation and inference, clinicians sometimes express skepticism that any structured assessment instrument could do justice to the richness, complexity, and uniqueness of human psychological functioning. However, SWAP statements can be combined in virtually infinite patterns to capture complex, nuanced psychological phenomena and convey meanings that transcend the content of the individual items. The whole is greater than the sum of its parts. (The mathematically inclined reader might consider that the number of possible orderings of SWAP statements vastly exceeds the earth's population.)

By way of illustration, consider the meaning of the SWAP item "Tends to be sexually seductive or provocative." If a patient receives a high score on this item along with high scores on the items "Has an exaggerated sense of self-importance (e.g., feels special, superior, grand, or envied)" and "Seems to treat others primarily as an audience to witness own importance, brilliance, beauty, etc.," a portrait begins to emerge of a narcissistically organized person who may seek sexual attention to bolster a sense of being special and uniquely desirable. If this same patient also receives high scores on the items "Tends to feel s/he is not his/her true self with others; may feel false or fraudulent" and "Tends to feel s/he is inadequate, inferior, or a failure," a more complex portrait begins to emerge. The items, in combination, indicate that grandiosity coexists with painful feelings of inadequacy and serves to mask or compensate for them. This duality lies at the heart of narcissistic personality pathology (Russ, Bradley, Shedler, & Westen, 2008).

Alternatively, if the SWAP item describing sexual seductiveness is combined with the items "Tends to fear s/he will be rejected or abandoned," "Appears to fear being alone; may go to great lengths to avoid being alone," and "Tends to be ingratiating or submissive (e.g., consents to things s/he does not want to do, in the hope of getting support or approval)," a portrait begins to emerge of a person with a dependent personality style, who may rely on sexuality as a desperate means of maintaining attachments in the face of feared abandonment.

If the sexual seductiveness item is instead combined with the items "Tends to act impulsively (e.g., acts without forethought or concern for consequences)," "Takes advantage of others; has little investment in moral values (e.g., puts own needs first, uses or exploits people with little regard for their feelings or welfare, etc.)," and "Experiences little or no remorse for harm or

injury caused to others,” a portrait begins to emerge of a person with a psychopathic personality style who seeks immediate need gratification and has no qualms about exploiting others sexually.

These examples illustrate how SWAP items can be combined to communicate complex clinical concepts and how a single SWAP item can convey a range of meanings depending on the items that surround and contextualize it. I will further illustrate this with a case example (see Bridging Diagnosis and Clinical Case Formulation, below).

Diagnosis, Syndromal and Dimensional

The SWAP-200 generates 37 diagnostic scale scores organized into three score profiles.³ The score profiles provide (a) dimensional scores for *DSM-IV* and *DSM-5* (American Psychiatric Association, 2013) personality disorder diagnoses, (b) dimensional scores for an alternative set of personality syndromes identified empirically through SWAP research (see An Improved Taxonomy for Personality Diagnosis, below), and (c) dimensional trait scores derived via factor analysis of the SWAP item set (Shedler & Westen, 2004a). The SWAP also generates a Psychological Health Index, which measures adaptive psychological resources and capacities, or ego strengths. The SWAP National Security Edition includes the Dispositional Indicators of Risk Exposure (DIRE) scale, which was developed in collaboration with agencies of the United States federal government to assess the potential for destructive or high-risk behavior in personnel employed in, or being evaluated for, sensitive positions such as those requiring access to classified information (Shechter & Lang, 2011).

SWAP diagnostic scores are expressed as T-scores ($M = 50$, $SD = 10$) and graphed to create score profiles, as shown in Figure 9.1.⁴ SWAP-200 personality disorder scales measure the similarity or “match” between a patient and diagnostic prototypes representing each *DSM* personality disorder in its pure or “ideal” form (e.g., a prototypical patient with paranoid personality disorder). Thus, personality disorders are assessed on a continuum: Low scores indicate that the patient does not resemble or match the diagnostic prototype, and high scores indicate a strong match.

Where categorical diagnosis is desired (e.g., to facilitate clinical communication or for “backward compatibility” with the categorical approach

³SWAP-200 is available for clinical use at the time of this writing (see Shedler, 2009). SWAP-II will be available in the future and will similarly provide score profiles for *DSM* diagnoses, empirically identified personality syndromes, and factor-analytically derived trait dimensions. Computational algorithms for SWAP-II differ from those of SWAP-200 (see Westen et al., 2012, 2014).

⁴For descriptions of scale construction methods for SWAP-200, see Shedler and Westen (2004b) and Westen and Shedler (1999a, 1999b). For descriptions of scale construction methods for SWAP-II, see Westen et al. (2012, 2014).

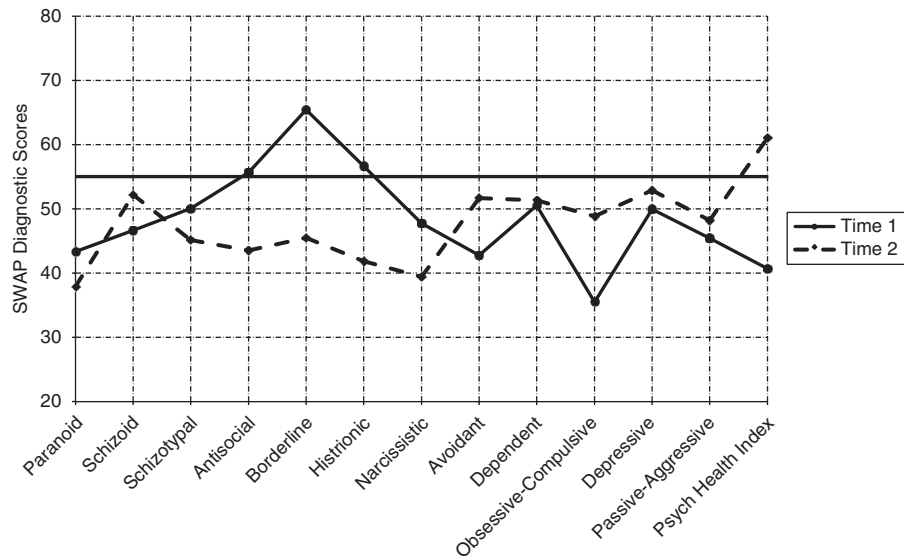


Figure 9.1. SWAP-200 Personality Disorder Score Profile (DSM-5 diagnoses).

of DSM), a cut-score of $T \geq 60$ provides a threshold for assigning a categorical diagnosis and a score of $T \geq 55$ warrants a diagnosis of traits or features of a PD. Thus, the patient represented by the solid line in Figure 9.1 would receive a DSM diagnosis of “borderline personality disorder with antisocial and histrionic traits.”

This approach to dimensional diagnosis preserves a *syndromal* understanding of personality. That is, it views personality as a configuration of functionally interrelated psychological processes (encompassing, for example, patterns of thinking, feeling, motivation, interpersonal functioning, coping, and defending). *Functionally related* means that the personality processes are interdependent, are causally linked, and form a coherent, recognizable configuration or whole (cf. Ahn, 1999; Cantor & Genero, 1986; Kim & Ahn, 2002).

Dimensional diagnosis follows from the recognition that all personality syndromes fall on a continuum from relatively healthy through severely disturbed. For example, a relatively healthy person with an obsessional personality style might be precise, orderly, logical, more comfortable with ideas than feelings, a bit more concerned than most with authority and control, and somewhat rigid in certain areas of thought and behavior. Toward the more disturbed end of the obsessional spectrum are individuals who are rigidly dogmatic. They have little conscious access to affect, are preoccupied with control, and misapply logic in ways that lead them to “miss the forest for the

trees.” Such individuals might properly be described as having a “disorder,” but the threshold for diagnosing a disorder is a cut-point on a continuum (similar to many diagnoses in medicine, where variables such as blood pressure are measured on a continuum but physicians refer to certain ranges as “borderline” or “high”).

Although I am emphasizing here the utility of a syndromal approach to personality, SWAP also provides dimensional trait scores, derived via factor analysis of the SWAP item set. Factor analysis of the SWAP item sets yields clinically and empirically coherent personality factors, 12 in the case of SWAP-200 (Shedler & Westen, 2004a) and 14 in the case of SWAP-II (Westen, Waller, Shedler, & Blagov, 2014). These dimensional trait or factor scores provide an additional source of information to supplement syndromal diagnosis.

Syndromal and trait models of personality serve different purposes. Among other things, the former is *person-centered* (focusing on kinds of people) and the latter is *variable-centered* (focusing on kinds of variables). Elsewhere, I have suggested that a diagnostic system is like a good map in that it must accurately depict the territory (Shedler et al., 2010). However, sometimes one requires a road map, sometimes a topographical map depicting elevations, and sometimes a political map. A road map, regardless of its validity, is of little use to a mountaineer, and a topographical map will be of little use to a motorist. One consequence of the science–practice schism is that there has been virtually no constructive discussion in the field about what kind of map is useful when. Instead, different constituencies have simply lobbied for maps that serve their own purposes, citing reliability and validity but failing to recognize that the wrong reliable and valid map can leave a different kind of traveler stranded on a cliff.

BRIDGING DIAGNOSIS AND CLINICAL CASE FORMULATION

Diagnosis and clinical case formulation are often viewed as separate activities. SWAP bridges these activities, allowing clinicians and investigators to both make psychiatric diagnoses and derive detailed clinical case formulations from the same item set.⁵ I will illustrate with a clinical case example.

⁵This section contains ideas first published in Lingardi, Shedler, and Gazzillo (2006). See the original publication for a more complete description of the case, treatment methods, and findings. The narrative description presented here incorporates some items from the revised SWAP-II and therefore differs slightly from the description in the original publication.

Background

“Melanie” is a 30-year-old woman with chief complaints of substance abuse and an inability to extricate herself from an abusive relationship. She was diagnosed with substance abuse, based on the Structured Clinical Interview for *DSM-IV* Axis I Disorders (SCID), and with Borderline Personality Disorder with histrionic traits, based on the SCID-II. She received a Global Assessment of Functioning (GAF) score of 45 at intake, indicating significant impairment in functioning.

Melanie’s early family environment was one of neglect and family strife. A recurring scenario is illustrative: Melanie’s mother would scream at her husband and say she was leaving him, then lock herself in her room, leaving Melanie frightened and in tears. Both parents would then ignore Melanie and often forget to feed her. By adolescence, Melanie was skipping school and spending her days sleeping or wandering the streets. At age 18 she left home and began “life on the streets,” entering a series of chaotic sexual relationships, abusing street drugs, and engaging in petty theft. In her mid-20s, she moved in with her boyfriend, a small-time drug dealer. She periodically prostituted herself to obtain money or drugs for him.

Melanie began psychodynamic therapy at a frequency of three sessions per week. The first 10 sessions were recorded and transcribed. Two clinicians, blind to other data, reviewed the transcripts and scored the SWAP-200 based on the session transcripts. The SWAP-200 scores were averaged across the two clinical judges to enhance reliability and obtain a single SWAP description. After 2 years of psychotherapy, 10 consecutive psychotherapy sessions were again recorded and transcribed and the SWAP evaluation was repeated.

Descriptive Diagnosis

The solid line in Figure 9.1 (Time 1) shows Melanie’s SWAP-200 scores profile for *DSM-IV/DSM-5* personality disorder diagnoses. Higher scale scores indicate more severe disturbance. The Psychological Health Index, which reflects clinicians’ consensual understanding of healthy personality functioning (Westen & Shedler, 1999a), is graphed as well. Higher scores on the Psychological Health Index indicate greater psychological strengths and resources.

Melanie’s score profile shows a marked elevation for borderline personality ($T = 65$, or one and a half standard deviations above the reference sample mean), with secondary elevations for histrionic personality disorder ($T = 57$) and antisocial personality disorder ($T = 56$). After application of the recommended cut-scores of $T \geq 60$ for making a categorical personality disorder diagnosis and $T \geq 55$ for diagnosing traits or features, Melanie’s *DSM* diagnosis at the start of treatment (Time 1) is “borderline personality disorder

with histrionic and antisocial traits.” Also noteworthy is the low T-score of 41 on the Psychological Health Index, nearly a standard deviation below the mean in a reference sample of patients with personality disorder diagnoses. The low score indicates significant psychological deficit and impairment.

Narrative Case Description

Clinicians can move from diagnosis to individualized case description by shifting focus from diagnostic scale scores to individual SWAP items. A narrative description can be created simply by selecting and listing the 30 SWAP items with the highest scores (i.e., those scored 5, 6, or 7) or (preferably) arranging them in paragraph form to create a narrative description.

The narrative description for Melanie, below, illustrates this approach. The description is constructed exclusively from the 30 SWAP items with scores of 5 or above. To aid the flow of the text, I have grouped conceptually related items, made minor grammatical edits, and added some topic sentences and connecting text (italicized).

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

Melanie has personality features associated specifically with borderline personality. 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. She often acts in such a way as to elicit her own feelings in other people (e.g., provoking anger when she herself is angry, inducing anxiety in others when she herself is anxious), and she tends to draw people into scenarios or pull them into roles that they experience as alien and unfamiliar (e.g., being uncharacteristically cruel, feeling like the only person in the world who can help).

When upset, Melanie has difficulty perceiving positive and negative qualities in the same person at the same time (e.g., she sees others in black or white terms and may shift suddenly from seeing someone as caring to seeing them as malevolent). She expresses contradictory feelings without being disturbed by the inconsistency and seems to have little need to reconcile or resolve contradictory ideas. 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 inside. *Her 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 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, Melanie's attitudes toward men and sexuality are problematic and conflictual. She tends to be hostile toward members of the opposite sex (whether consciously or unconsciously), and she associates sexual activity with danger (e.g., injury or punishment). She appears afraid of commitment to a long-term love relationship, instead choosing partners who seem inappropriate in terms of age, status (e.g., social, economic, intellectual), or other factors.

This narrative case description provides an in-depth portrait of a troubled patient with borderline personality pathology, highlighting personality features such as splitting, projective identification, identity diffusion, and affect dysregulation. The description illustrates the difference between descriptive psychiatry (aimed at establishing a diagnosis) and clinical case formulation (aimed at understanding the psychological makeup of an individual person). However, all the findings presented here are derived from the same quantitative SWAP data.

Melanie's case has a happy ending. The dashed line in Figure 9.1 shows Melanie's PD scores after 2 years of psychotherapy (Time 2). Her scores on the Borderline, Histrionic, and Antisocial dimensions have dropped below $T = 50$, and she no longer warrants any DSM PD diagnosis. Her score on the Psychological Health Index has increased by two standard deviations, from 41 to 61, indicating the development of substantial psychological resources and capacities (for a more complete discussion of this case, see Lingardi, Shedler, & Gazzillo, 2006).

Reliability and Validity

Interrater reliability of SWAP diagnostic scale scores is above .80 in all studies to date and is often above .90 (Marin-Avellan et al., 2005; Westen & Muderrisoglu, 2003; Westen & Shedler, 2007). Median test-retest reliability

of SWAP-II personality disorder scales, over a 4- to 6-month time interval, is .90, with a range of .86 to .96 for individual scales. Median test–retest reliability for SWAP-II factor (dimensional trait) scales is .85, with a range of .77 to .96 (Blagov et al., 2012). Median alpha reliability (Cronbach’s alpha) for diagnostic scales for SWAP-II empirically derived personality syndromes (see An Improved Taxonomy for Personality Diagnosis, below) is .79, with a range from .72 to .94. These are strong reliability findings that compare favorably with reliabilities for structured interviews and questionnaires that minimize or eliminate clinical inference.

With respect to validity, SWAP diagnostic scales show predicted relations with an extensive range of external criterion variables in both adult and adolescent samples. These include genetic history variables such as psychosis and substance abuse in first- and second-degree biological relatives; developmental history variables such as childhood physical abuse, sexual abuse, animal torture, fire setting, truancy, and other school-related problems; life events such as psychiatric hospitalizations, suicide attempts, arrests, violent criminal behavior, and perpetrating domestic abuse; ratings of occupational functioning, social functioning, and global adaptive functioning; response to mental health treatment, and numerous other measures (see, e.g., Marin-Avellan et al., 2005; Shedler & Westen, 2004a; Westen & Muderrisoglu, 2003, 2006; Westen & Shedler, 1999a, 1999b, 2007; Westen, Shedler, Durrett, Glass, & Martens, 2003; Westen & Weinberger, 2004). For reviews, see Westen and Shedler (2007) and Blagov et al. (2012).

There is a well-established literature on the limitations of clinical judgment, and it is fair to ask why the SWAP yields strong reliability and validity findings that seem incongruous with this literature. The answers are straightforward. First, studies of clinical judgment have too often asked clinicians to make predictions about things that fall well outside their legitimate expertise (Westen & Weinberger, 2004).⁶ In contrast, the SWAP does not ask clinicians to *predict* anything, only to describe patients they know, based on psychological information that is readily available to them. Second, studies of clinical judgment often do not use appropriate psychometric methods to quantify clinical judgment in a reliable way, as the SWAP is designed to do. Third, studies of clinical judgment typically conflate clinicians’ ability to provide accurate information about their patients (which they do well) with their ability to combine and weight variables to make predictions (a task necessarily performed better by statistical methods). In fact, a substantial literature documents the reliability and validity of clinical observation and inference when it is quantified and utilized appropriately (see Westen

⁶Unfortunately, some clinicians have been all too willing to offer such prognostications.

& Weinberger, 2004). It is unfortunate (and telling) that research on the limitations of clinical judgment is widely cited in academic psychology while compelling research on its strengths is often overlooked.

The SWAP differs from other assessment approaches in that it harnesses clinical judgment with psychometric methods designed for this purpose, then applies statistical and actuarial methods to the resulting quantitative data. In short, it relies on clinicians to do what they do best: describe individual patients they know well. It relies on statistical algorithms to do what they do best: combine data optimally to derive reliable and valid scales and maximize prediction. In the framework of Paul Meehl's (1954) classic text *Clinical Versus Statistical Prediction*, the SWAP would be considered an example of *statistical prediction*.

AN IMPROVED TAXONOMY FOR PERSONALITY DIAGNOSIS

The system for personality diagnosis provided by the *DSM* finds little favor with clinicians or researchers (Shedler & Westen, 2004b, 2007; Westen & Shedler, 1999a). The *DSM-5* Personality and Personality Disorders Work Group attempted to replace it entirely, but ideological disputes prevented the work group from producing a viable alternative. As a result, *DSM-5* defaulted to *DSM-IV* diagnostic categories and criteria, and the opportunity for an improved, officially sanctioned system for personality diagnosis was lost.

An optimal diagnostic system should (a) “carve nature at the joints” as closely as nature reveals them and available research methods permit; (b) provide descriptions of personality syndromes that are clinically useful and relevant—ideally, they should facilitate a level of understanding that can guide treatment decisions; and (c) provide a sound, workable method for making diagnoses in day-to-day clinical practice.

An alternative to developing a diagnostic system by committee (with the unavoidable influences of group dynamics, politics, ideology, and other biases) is to derive a diagnostic taxonomy empirically. Doing so entails conducting comprehensive assessments of personality in large, clinically representative patient samples, then employing statistical methods to identify and describe naturally occurring diagnostic groupings, assuming such groupings exist.

My coinvestigators and I first described a diagnostic system based on such an approach in 1999, identifying naturally occurring diagnostic groupings in a national sample of personality disordered patients assessed with the SWAP-200 (Westen & Shedler, 1999b). In this section, I summarize the findings of newer research using the SWAP-II in a larger, more representative sample ($N = 1,201$) of adult patients (Westen et al., 2012). We used the method of Q-factor analysis to identify naturally occurring diagnostic

groupings. Q-factor analysis is computationally identical to factor analysis, with the difference that factor analysis identifies groupings of similar *variables*, whereas Q-factor analysis identifies groupings of similar *cases* or *people*. The resulting diagnostic groupings are data-driven and are not the product of a committee decision process.

Data were provided by 1,201 licensed psychologists or psychiatrists, each of whom used the SWAP-II to describe a single, randomly selected current patient. The clinicians were instructed to describe “an adult patient you are currently treating or evaluating who has enduring patterns of thoughts, feelings, motivation, or behavior—that is, personality patterns—that cause distress or dysfunction.” To ensure a clinically representative sample, the instructions emphasized that patients need not have a DSM PD diagnosis. The methods are described in our original research report (Westen et al., 2012).

An Empirically Derived Personality Taxonomy

The research identified 10 distinct, empirically and clinically coherent personality syndromes (Q-factors) organized hierarchically under superordinate groupings or broad personality spectra. Figure 9.2 illustrates the hierarchical structure of the empirically derived diagnostic system. At the level of broad superordinate groupings, the analysis identified an internalizing spectrum of personality syndromes, an externalizing spectrum, a borderline-dysregulated spectrum, and a spectrum labeled neurotic styles.

Individuals with syndromes in the *internalizing* spectrum experience chronic painful emotions, especially depression and anxiety. They tend to be emotionally constricted and socially avoidant and tend to blame themselves for their difficulties. The spectrum subsumes the diagnoses of Depressive Personality, Anxious-Avoidant Personality, Dependent-Victimized Personality, and Schizoid-Schizotypal Personality.

Individuals with syndromes in the *externalizing* spectrum are angry or hostile, self-centered, and lacking in empathy. They blame others for their difficulties. The spectrum subsumes the diagnoses of Antisocial-Psychopathic Personality, Narcissistic Personality, and Paranoid Personality.

Individuals in the *borderline-dysregulated* spectrum are qualitatively distinct from stable internalizers or stable externalizers. Their perceptions of self and others are unstable and changeable, and they show impaired ability to regulate emotion. As a result, they tend to oscillate between emotions characteristic of both internalizing and externalizing spectrum pathology (e.g., depression, anxiety, rage). They may best be described as “stably unstable” (Schmideberg, 1959). The salience of affect dysregulation in the clinical

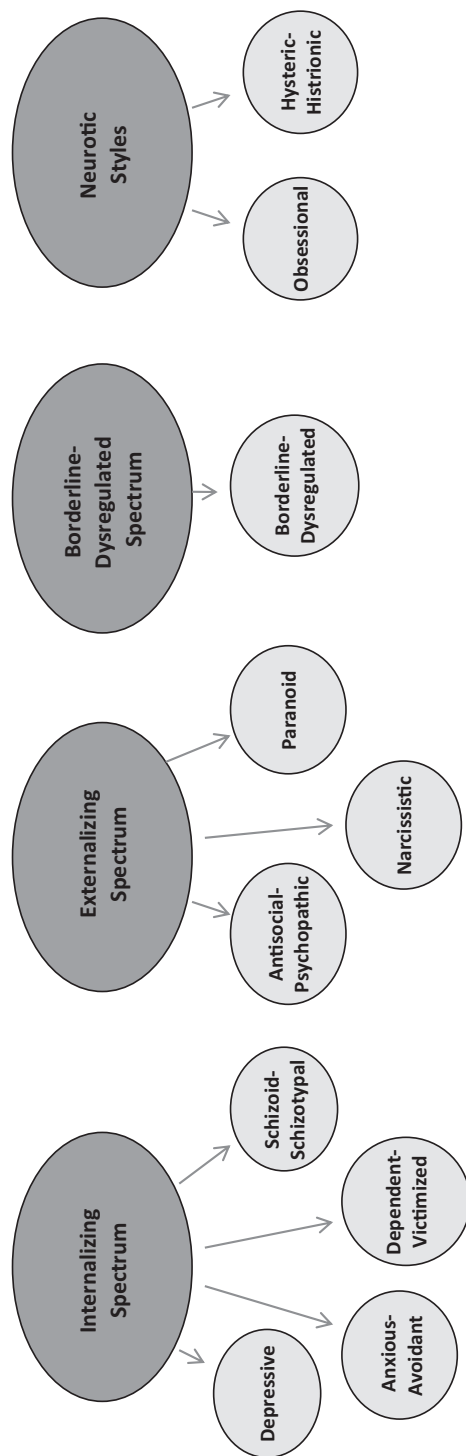


Figure 9.2. Hierarchical structure of personality diagnoses.

picture led my coinvestigators and I to hyphenate the name of the syndrome and add *dysregulated* to the more familiar term *borderline*.

The *neurotic styles* spectrum subsumes the diagnoses of Obsessional Personality and Hysterical-Histrionic Personality. The name of the spectrum reflects the recognition that individuals with these personality syndromes are, on average, higher functioning than those in the other diagnostic groupings and may or may not show a level of impairment that warrants the term *disorder*. The two personality syndromes resemble the character styles or “neurotic styles” described in the clinical literature (e.g., MacKinnon, Michels, & Buckley, 2009; McWilliams, 2011; Shapiro, 1965) more than they resemble DSM descriptions of obsessive–compulsive and histrionic personality disorders. The framers of *DSM–III* amplified the level of pathology of these personality syndromes to fit them into a medical-model taxonomy of disorders. Unfortunately, the resulting DSM diagnostic criteria described caricatures, not the patients generally seen in real-world practice.

Empirically Derived Descriptions of Personality Syndromes

In addition to identifying naturally occurring personality syndromes, our research method allowed us to generate an empirically-derived description of each personality syndrome. A description of the core, defining features of each diagnostic grouping or syndrome was obtained simply by listing the SWAP items with the highest factor scores for the syndrome. I will use borderline-dysregulated personality for illustration.

Exhibit 9.1 lists the 24 SWAP items with the highest factor scores for borderline-dysregulated personality (the items most central to the syndrome).⁷ A number of findings are noteworthy. First, the emergence of this diagnostic grouping in the Q-factor analysis validates the concept of borderline personality as a diagnostic entity. Second, the items describe a psychologically richer and more complex syndrome than described by the DSM. Third, the description addresses internal psychological processes and aspects of inner experience that are crucial to understanding and treating this personality syndrome.

The findings also indicate that clinical theories that view splitting, projective identification, and related psychological processes as central to borderline personality are, in fact, accurate. Overall, the empirically-derived personality syndrome more closely resembles the concept of borderline

⁷To facilitate understanding of this complex syndrome, I have grouped the items under several broad themes. Another investigator might quibble with my method of organization, but the headings and item groupings serve only to facilitate presentation; they have no bearing on the definition or assessment of the personality syndrome.

EXHIBIT 9.1
Empirically-Derived Description of Borderline-Dysregulated Personality

Affect dysregulation

Emotions tend to change rapidly and unpredictably.
Emotions tend to spiral out of control, leading to extremes of anxiety, sadness, rage, etc.
Tends to become irrational when strong emotions are stirred up; may show a significant decline from customary level of functioning.
Is prone to intense anger, out of proportion to the situation at hand (e.g., has episodes of rage).
Is unable to soothe or comfort him/herself without the help of another person (i.e., has difficulty regulating own emotions).
Tends to “catastrophize”; is prone to see problems as disastrous, unsolvable, etc.
Tends to feel unhappy, depressed, or despondent.

Splitting

When upset, has trouble perceiving both positive and negative qualities in the same person at the same time; sees others in black or white terms (e.g., may swing from seeing someone as caring to seeing him/her as malevolent and intentionally hurtful).
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).

Projective identification

Manages to elicit in others feelings similar to those she/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).
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).

Identity diffusion

Lacks a stable sense of who she/he is (e.g., attitudes, values, goals, and feelings about self seem unstable or ever-changing).
Is prone to painful feelings of emptiness (e.g., may feel lost, bereft, abjectly alone even in the presence of others).

Insecure attachment

Tends to be needy or dependent.
Appears to fear being alone; may go to great lengths to avoid being alone.
Tends to fear she/he will be rejected or abandoned.
Tends to become attached quickly or intensely; develops feelings, expectations, etc. that are not warranted by the history or context of the relationship.
Tends to feel misunderstood, mistreated, or victimized.

Self-harm (desperate efforts to self-regulate)

Tends to engage in self-mutilating behavior (e.g., self-cutting, self-burning).
Tends to make repeated suicidal threats or gestures, either as a “cry for help” or as an effort to manipulate others.
Struggles with genuine wishes to kill him/herself.

Chaotic lifestyle

Relationships tend to be unstable, chaotic, and rapidly changing.
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).
Tends to be impulsive.

personality *organization* described in the clinical literature (e.g., Clarkin et al., 2006; Kernberg, 1975, 1984; McWilliams, 2011) than it does the DSM description of borderline personality disorder.

The items or personality features constituting the description of borderline-dysregulated personality (and all of the other empirically-identified syndromes) cannot be explained away as artifacts of clinicians' theoretical preconceptions. They emerged repeatedly when we stratified the sample by the theoretical orientation of the reporting clinicians, with the same items ranked highly by psychodynamic clinicians, cognitive-behavioral clinicians, humanistic clinicians, biologically-oriented clinicians, and so on.

Psychometric Assessment With the SWAP

We developed the SWAP-II diagnostic scales to assess the empirically derived diagnostic syndromes by summing the most descriptive SWAP-II items for each syndrome (thus, the diagnostic scale for borderline-dysregulated personality comprises the 24 items listed in Exhibit 9.1). The number of scale items ranges from a low of 14 (for paranoid personality) to a high of 24 (for borderline-dysregulated personality), with the number of items reflecting the psychological complexity of the syndrome. Alpha reliabilities for the diagnostic scales range from .72 to .94, with a median reliability of .79. All diagnostic scores are scaled as normalized T-scores ($M = 50$, $SD = 10$) to facilitate test interpretation.

An empirically derived Psychological Health Index was created by the same method and yielded an additional scale assessing global personality health/dysfunction. All personality syndromes fall on a continuum of functioning, and the score on the Psychological Health Index provides a context for interpreting other SWAP scale scores. An elevated score for a personality syndrome, coupled with a high Psychological Health Index score, indicates that the person is functioning at the healthier end of the health-pathology continuum for that syndrome, and a low score on the Psychological Health Index indicates the opposite. For example, a patient with an elevated score for paranoid personality and a high Psychological Health Index score has meaningful psychological resources (ego strengths) to draw upon and may be able to make constructive use of psychotherapy. A patient with the same paranoid personality score and a low Psychological Health Index score may prove untreatable. Both patients are likely to incorporate the therapist into a paranoid worldview, suspecting the therapist of nefarious motives and intentions. However, the first patient will likely retain a capacity to reflect on his experience of the therapist and call his perceptions into question, whereas the second patient may not.

Diagnosis in Day-to-Day Practice

When maximum psychometric precision is required or where there are challenging diagnostic dilemmas, assessors can describe patients with the SWAP and obtain diagnostic scale scores for all the empirically-derived personality syndromes (as well as for *DSM* personality disorder diagnoses and for SWAP factors or trait dimensions). For day-to-day diagnosis, my coinvestigators and I have proposed a diagnostic system based on *prototype matching*.

In prototype matching diagnosis, the descriptions of the empirically-derived personality syndromes are presented in paragraph rather than list form, to create a narrative description of each personality syndrome. The narrative descriptions constitute *diagnostic prototypes* that describe each syndrome in its “ideal” or “pure” form. The diagnostic prototypes comprise the SWAP-II items that are most defining of each syndrome (the same items used to construct the psychometric scales), organized and edited to create narratively coherent paragraphs. Each prototype description is preceded by a single-sentence summary intended to orient the diagnostician and convey telegraphically the core features of the syndrome.

The diagnostician’s task is to consider the prototype description as a whole—that is, as a configuration or gestalt—and rate the overall similarity or match between a patient and the diagnostic prototype. The resulting diagnosis is dimensional (a 1–5 rating), but the scale can be dichotomized when a categorical (present/absent) diagnosis is desired (with ratings ≥ 4 indicating “caseness”).

Figure 9.3 illustrates the prototype matching approach to personality diagnosis, with depressive personality used as an example. Despite its omission from the *DSM*, depressive personality appears to be the most prevalent personality syndrome seen in the community (Westen & Shedler, 1999b). Diagnostic prototypes for all of the empirically derived personality syndromes are reproduced in our original research report (Westen et al., 2012) and are available for download from <http://www.SWAPassessment.org>.

Prototype matching works *with*, rather than against, the naturally occurring cognitive decision processes of diagnosticians. It has considerable advantages over the criterion-counting approach of the *DSM*, leading to improved diagnostic reliability and validity as well as reduced diagnostic comorbidity. In head-to-head comparisons clinicians rated the SWAP prototype matching approach as more clinically useful and relevant than both the *DSM* diagnostic system and dimensional trait models of personality (e.g., Rottman et al., 2009; Spitzer et al., 2008). The conceptual rationale for the prototype matching method and the research evidence supporting it are described in detail elsewhere (see Westen, DeFife, Bradley, & Hilsenroth, 2010; see also Ortigo,

Summary statement: Individuals With Depressive Personality are prone to feelings of depression and inadequacy, tend to be self-critical or self-punitive, and may be preoccupied with concerns about abandonment or loss.

Individuals who match this prototype tend to feel depressed or despondent and to feel inadequate, inferior, or a failure. They tend to find little pleasure or satisfaction in life's activities and to feel life has no meaning. They are insufficiently concerned with meeting their own needs, disavowing or squelching their hopes and desires to protect against disappointment. They appear conflicted about experiencing pleasure, inhibiting feelings of excitement, joy, or pride. They may likewise be conflicted or inhibited about achievement or success (e.g., failing to reach their potential or sabotaging themselves when success is at hand). Individuals who match this prototype are generally self-critical, holding themselves to unrealistic standards and feeling guilty and blaming themselves for bad things that happen. They appear to want to "punish" themselves by creating situations that lead to unhappiness or avoiding opportunities for pleasure and gratification. They have trouble acknowledging or expressing anger and instead become depressed, self-critical, or self-punitive. Individuals who match this prototype often fear that they will be rejected or abandoned, are prone to painful feeling of emptiness, and may feel bereft or abjectly alone even in the presence of others. They may have a pervasive sense that someone or something necessary for happiness has been lost forever (e.g., a relationship, youth, beauty, success).

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

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

Figure 9.3. Depressive Personality Prototype.

Bradley, & Westen, 2010; Shedler & Westen, 2004b; Westen, Shedler, & Bradley, 2006).

CONCLUSION: INTEGRATING CLINICAL AND EMPIRICAL PERSPECTIVES

A clinically useful diagnostic system should encompass the spectrum of personality syndromes seen in clinical practice and have meaningful treatment implications. An empirically sound diagnostic system should facilitate reliable and valid diagnoses: Independent clinicians should be able to

arrive at the same diagnosis, diagnoses should be distinct, and each diagnosis should be associated with conceptually meaningful correlates, antecedents, and sequelae.

An obstacle to achieving this ideal has been the persistent schism between science and practice in the mental health professions. Too often, empirical research has been conducted in isolation from the crucial data of clinical observation, and clinical theory has developed without regard for empirical credibility. Professional researchers and professional clinicians tend to talk past rather than with one another.

SWAP research represents an effort to bridge the science–practice schism by quantifying clinical observation and expertise, thus making clinical constructs accessible to empirical study. It relies on clinicians to make observations and inferences about individual patients they know and relies on quantitative methods to reveal relationships and combine data in optimal ways.

The SWAP provides a “language” for clinical case description that is both psychometrically sound and clinically rich enough to describe the complexities of real patients. A sizable schism remains between science and practice. Perhaps the SWAP will provide a language all parties can speak.

REFERENCES

- Ahn, W.-K. (1999). Effect of causal structure on category construction. *Memory & Cognition*, *27*, 1008–1023. doi:10.3758/BF03201231
- American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.
- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: Author.
- Betan, E., Heim, A. K., Conklin, C. Z., & Westen, D. (2005). Countertransference phenomena and personality pathology in clinical practice: An empirical investigation. *The American Journal of Psychiatry*, *162*, 890–898. doi:10.1176/appi.ajp.162.5.890
- Blagov, P. S., Bi, W., Shedler, J., & Westen, D. (2012). The Shedler–Westen Assessment Procedure (SWAP): Evaluating psychometric questions about its reliability, validity, and impact of its fixed score distribution. *Assessment*, *19*, 370–382. doi:10.1177/1073191112436667
- Block, J. (1971). *Lives through time*. Berkeley, CA: Bancroft.
- Block, J. (1978). *The Q-sort method in personality assessment and psychiatric research*. Palo Alto, CA: Consulting Psychologists Press. (Original work published 1961)

- Cantor, N., & Genero, N. (1986). Psychiatric diagnosis and natural categorization: A close analogy. In T. Millon & G. L. Klerman (Eds.), *Contemporary directions in psychopathology: Toward the DSM-IV* (pp. 233–256). New York, NY: Guilford Press.
- Clarkin, J. F., Yeomans, F. E., & Kernberg, O. F. (2006). *Psychotherapy for borderline personality: Focusing on object relations*. Washington, DC: American Psychiatric Publishing.
- Cousineau, T. M., & Shedler, J. (2006). Predicting physical health: Implicit mental health measures versus self-report scales. *Journal of Nervous and Mental Disease*, *194*, 427–432. doi:10.1097/01.nmd.0000221373.74045.51
- Gabbard, G. O. (2005). *Psychodynamic psychiatry in clinical practice* (4th ed.). Washington, DC: American Psychiatric Publishing.
- 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, NY: Guilford Press.
- Kernberg, O. (1975). *Borderline conditions and pathological narcissism*. New York, NY: Aronson.
- Kernberg, O. (1984). *Severe personality disorders*. New Haven, CT: Yale University Press.
- Kernberg, O. (1996). A psychoanalytic theory of personality disorders. In J. F. Clarkin & M. F. Lenzenweger (Eds.), *Major theories of personality disorder* (pp. 106–137). New York, NY: Guilford 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: General*, *131*, 451–476. doi:10.1037/0096-3445.131.4.451
- Kohut, H. (1971). *The analysis of the self*. New York, NY: International Universities Press.
- Linehan, M. M. (1993). *Cognitive-behavioral treatment of borderline personality disorder*. New York, NY: Guilford Press.
- Lingiardi, V., Shedler, J., & Gazzillo, F. (2006). Assessing personality change in psychotherapy with the SWAP-200: A case study. *Journal of Personality Assessment*, *86*, 23–32. doi:10.1207/s15327752jpa8601_04
- Livesley, W. J. (Ed.). (1995). *The DSM-IV personality disorders*. New York, NY: Guilford Press.
- MacKinnon, R., Michels, R., & Buckley, P. (2009). *The psychiatric interview in clinical practice* (2nd ed.). Washington, DC: American Psychiatric Publishing.
- Marin-Avellan, L. E., McGauley, G., Campbell, C., & Fonagy, P. (2005). Using the SWAP-200 in a personality-disordered forensic population: Is it valid, reliable and useful? *Journal of Criminal Behaviour and Mental Health*, *15*, 28–45. doi:10.1002/cbm.35
- McCrae, R., & Costa, P., Jr. (1990). *Personality in adulthood*. New York, NY: Guilford Press.

- McWilliams, N. (2011). *Psychoanalytic diagnosis: Understanding personality structure in the clinical process* (2nd ed.). New York, NY: Guilford Press.
- Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence*. doi:10.1037/11281-000
- Meehl, P. E. (1973). Why I do not attend case conferences. In *Psychodiagnosis: Selected papers* (pp. 225–302). Minneapolis: University of Minnesota Press.
- Ortigo, K. M., Bradley, B., & Westen, D. (2010). An empirically based prototype diagnostic system for DSM–V and ICD–11. In T. Millon, R. Krueger, & E. Simonsen (Eds.), *Contemporary directions in psychopathology: Scientific foundations of the DSM–V and ICD–11* (pp. 374–390). New York, NY: Guilford Press.
- Peebles, M. J. (2012). *Beginnings: The art and science of planning psychotherapy* (2nd ed.). New York, NY: Routledge.
- Perry, J. C., & Cooper, S. H. (1987). Empirical studies of psychological defense mechanisms. In R. Michels & J. O. Cavenar (Eds.), *Psychiatry* (Vol. 1, pp. 1–19). Philadelphia, PA: Lippincott.
- Reichenbach, H. (1938). *Experience and prediction*. Chicago, IL: University of Chicago Press.
- Rottman, B. M., Ahn, W. K., Sanislow, C. A., & Kim, N. S. (2009). Can clinicians recognize DSM–IV personality disorders from five-factor model descriptions of patient cases? *The American Journal of Psychiatry*, 166, 427–433. doi:10.1176/appi.ajp.2008.08070972
- Russ, E., Shedler, J., Bradley, R., & Westen, D. (2008). Refining the construct of narcissistic personality disorder: Diagnostic criteria and subtypes. *The American Journal of Psychiatry*, 165, 1473–1481. doi:10.1176/appi.ajp.2008.07030376
- Sawyer, J. (1966). Measurement and prediction, clinical and statistical. *Psychological Bulletin*, 66, 178–200. doi:10.1037/h0023624
- Schmideberg, M. (1959). The borderline patient. In S. Arieti (Ed.), *American handbook of psychiatry* (Vol. 1, pp. 398–416). New York, NY: Basic Books.
- Shapiro, D. (1965). *Neurotic styles*. New York, NY: Basic Books.
- Shechter, O. G., & Lang, E. L. (2011). *Identifying personality disorders that are security risks: Field test results* (Tech. Rep. No. 11-05). Retrieved from <http://www.dhramil/perserec/reports/tr11-05.pdf>
- Shedler, J. (2009). *Guide to SWAP-200 interpretation*. Retrieved from <http://www.SWAPassessment.org>
- Shedler, J., Beck, A. T., Fonagy, P., Gabbard, G. O., Gunderson, J., Kernberg, O., . . . Westen, D. (2010). Commentary: Personality disorders in DSM–5. *The American Journal of Psychiatry*, 167, 1026–1028. doi:10.1176/appi.ajp.2010.10050746
- Shedler, J., Mayman, M., & Manis, M. (1993). The illusion of mental health. *American Psychologist*, 48, 1117–1131. doi:10.1037/0003-066X.48.11.1117
- Shedler, J., & Westen, D. (1998). Refining the measurement of Axis II: A Q-sort procedure for assessing personality pathology. *Assessment*, 5, 333–353.

- Shedler, J., & Westen, D. (2004a). Dimensions of personality pathology: An alternative to the five-factor model. *The American Journal of Psychiatry*, *161*, 1743–1754.
- Shedler, J., & Westen, D. (2004b). Refining DSM–IV personality disorder diagnosis: Integrating science and practice. *The American Journal of Psychiatry*, *161*, 1350–1365.
- Shedler, J., & Westen, D. (2006). Personality diagnosis with the Shedler–Westen Assessment Procedure (SWAP): Bridging the gulf between science and practice. In Alliance Task Force (Ed.), *Psychodynamic diagnostic manual* (pp. 573–613). Silver Spring, MD: Alliance of Psychoanalytic Organizations.
- Shedler, J., & Westen, D. (2007). The Shedler–Westen Assessment Procedure (SWAP): Making personality diagnosis clinically meaningful. *Journal of Personality Assessment*, *89*, 41–55.
- Spitzer, R. L., First, M. B., Shedler, J., Westen, D., & Skodal, A. E. (2008). Clinical utility of five dimensional systems for personality diagnosis: A “consumer preference” study. *Journal of Nervous and Mental Disease*, *196*, 3567–374. doi:10.1097/NMD.0b013e3181710950
- Vaillant, G. (Ed.). (1992). *Ego mechanisms of defense: A guide for clinicians and researchers*. Washington, DC: American Psychiatric Press.
- Westen, D. (1991). Social cognition and object relations. *Psychological Bulletin*, *109*, 429–455. doi:10.1037/0033-2909.109.3.429
- Westen, D. (2004). *Clinical Diagnostic Interview*. Atlanta, GA: Emory University, Departments of Psychology and Psychiatry and Behavioral Sciences.
- Westen, D., DeFife, J. A., Bradley, B., & Hilsenroth, M. J. (2010). Prototype personality diagnosis in clinical practice: A viable alternative for DSM–5 and ICD–11. *Professional Psychology: Research and Practice*, *41*, 6, 482–487. doi:10.1037/a0021555
- Westen, D., Dutra, L., & Shedler, J. (2005). Assessing adolescent personality pathology: Quantifying clinical judgment. *British Journal of Psychiatry*, *186*, 227–238. doi:10.1192/bjp.186.3.227
- Westen, D., 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*, *2*, 355–364. doi:10.1037/1040-3590.2.4.355
- Westen, D., & Muderrisoglu, S. (2003). Reliability and validity of personality disorder assessment using a systematic clinical interview: Evaluating an alternative to structured interviews. *Journal of Personality Disorders*, *17*, 350–368.
- Westen, D., & Muderrisoglu, S. (2006). Clinical assessment of pathological personality traits. *The American Journal of Psychiatry*, *163*, 1285–1297. doi:10.1076/appi.ajp.163.7.1285
- Westen, D., Muderrisoglu, S., Fowler, C., Shedler, J., & Koren, D. (1997). Affect regulation and affective experience: Individual differences, group differences, and measurement using a Q-sort procedure. *Journal of Consulting and Clinical Psychology*, *65*, 429–439. doi:10.1037/0022-006X.65.3.429

- Westen, D., & Shedler, J. (1999a). Revising and assessing Axis II: I. Developing a clinically and empirically valid assessment method. *The American Journal of Psychiatry*, *156*, 258–272.
- Westen, D., & Shedler, J. (1999b). Revising and assessing Axis II: II. Toward an empirically based and clinically useful classification of personality disorders. *The American Journal of Psychiatry*, *156*, 258–272.
- Westen, D., & Shedler, J. (2007). Personality diagnosis with the Shedler–Westen Assessment Procedure (SWAP): Integrating clinical and statistical measurement and prediction. *Journal of Abnormal Psychology*, *116*, 810–822. doi:10.1037/0021-843X.116.4.810
- Westen, D., Shedler, J., Bradley, B., & DeFife, J. (2012). An empirically derived taxonomy for personality diagnosis: Bridging science and practice in conceptualizing personality. *The American Journal of Psychiatry*, *169*, 273–284. doi:10.1176/appi.ajp.2011.11020274
- Westen, D., Shedler, J., & Bradley, R. (2006). A prototype approach to personality diagnosis. *The American Journal of Psychiatry*, *163*, 846–856. doi:10.1176/appi.ajp.163.5.846
- 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. *The American Journal of Psychiatry*, *160*, 952–966. doi:10.1176/appi.ajp.160.5.952
- Westen, D., Waller, N. G., Shedler, J., & Blagov, P. S. (2014). Dimensions of personality and personality pathology: Factor structure of the Shedler–Westen Assessment Procedure–II (SWAP-II). *Journal of Personality Disorders*, *28*, 281–318. doi:10.1521/pedi_2012_26_059
- Westen, D., & Weinberger, J. (2004). When clinical description becomes statistical prediction. *American Psychologist*, *59*, 595–613. doi:10.1037/0003-066X.59.7.595
- Widiger, T. A., & Simonsen, E. S. (2005). Alternative dimensional models of personality disorders: Finding common ground. *Journal of Personality Disorders*, *19*, 110–130. doi:10.1521/pedi.19.2.110.62628