

A NEW LANGUAGE FOR PSYCHOANALYTIC DIAGNOSIS

Many psychoanalysts believe it is impossible to conduct empirical research without eviscerating or trivializing psychoanalytic constructs, and past research efforts have all too often reinforced this view. A new method for studying personality and personality pathology is presented that challenges such beliefs. This method, the Shedler-Westen Assessment Procedure (SWAP), captures the richness and complexity of psychoanalytic constructs and formulations while also providing reliable data for research. The method is being used to develop a new personality disorder taxonomy, as an alternative to the DSM, that is both empirically grounded and psychoanalytically relevant. Its role in psychoanalytic training and supervision is discussed, as is its value as a measure of structural change in psychoanalytic process and outcome research.

As for academic psychology, its methods give extensive and reliable data which are, in everybody's subjective assessment, of minor importance.

—CHARLES BRENNER

There sometimes seems to be an unbridgeable gulf between psychoanalysts and academic researchers. Academic research often seems superficial to psychoanalytic clinicians. The epigraph above simply expresses, in blunt form, the unspoken sentiment of many analysts. Researchers, for their part, often regard psychoanalytic formulations as unfalsifiable and unscientific. Perhaps there is merit to both lines of criticism.

This article discusses one effort to bridge the gulf. I will describe a method for studying personality and personality pathology that strives

Associate Professor, Graduate School of Professional Psychology, University of Denver.

Preparation of this paper was supported in part by NIMH grant MH60892, "Classification and Measurement of Personality Disorders." Submitted for publication February 9, 2001.

to capture the richness and complexity of psychoanalytic constructs and formulations without forsaking the benefits of empirical rigor. Before describing the work, I will first address some reservations analysts may harbor toward research.

SOME COMMENTS ON ANALYTIC SKEPTICISM

Analytic objections to research appear to fall into three categories: it is impossible, it is unnecessary, and it is irrelevant.

The first objection, that analytically meaningful research is impossible, may be the most prevalent view. Those who hold it might in principle welcome empirical research, but have been so disappointed by past research efforts that they have abandoned hope. They see no way to quantify important psychoanalytic constructs without trivializing and distorting them. They are turned off by the kind of reductionism that has characterized empirical research in the past. From the observation “it has not been done” they have reasoned “it cannot be done.” This paper is addressed primarily to these colleagues, in the hope of rekindling the idea that maybe it can be done after all.

430

The second objection, that empirical research is unnecessary, comes from analysts who believe the analytic method, by itself, is both necessary and sufficient for generating, testing, and revising analytic theory. This view has a long history. Brenner (1982) asserted that the analytic method is the “only” method that can shed light on the important aspects of mental life, and that evidence obtained by a single analyst behind closed doors is both “independent” and “objective.” Pine (1998) referred to the “experimentally ‘pure’ situation of psychoanalysis” and compared the analytic method to the microscope of the biologist and the telescope of the astronomer. The implication is that the analytic method can answer all relevant questions.

One problem with this position is that decades of “experimentally pure” analytic investigation have failed to resolve the theoretical differences that divide our field. We have competing models with different implications for technique, and advocates of each model regularly publish case studies “demonstrating” the superiority of their own approach. The case studies do not seem to change anyone’s mind, because analysts of different persuasions see different things in much the same case material (Gedo 1980). Ultimately, theoretical beliefs seem determined more by identifications with one’s own analyst, supervisors, and

institute than by weight of evidence. Another problem is that psychoanalytic case evidence has helped us demonstrate little to the satisfaction of anyone but other analysts. Perhaps the relative isolation of psychoanalysis in freestanding institutes has contributed to a certain innocence within psychoanalysis regarding evidential standards expected in other disciplines (on “the epistemic status of clinical data,” see Eagle 1987). Certainly health insurers, biological psychiatry, and advocates of “empirically validated therapies” (Chambless 1996; Task Force 1995) have found our evidence less than compelling.

The third objection, that empirical research is irrelevant, is held by analysts who think the rules of empirical science need not apply to psychoanalysis. One version of this objection is that psychoanalysis is hermeneutics, not science. Rather than resurrect hoary arguments, I will simply note that we need not succumb to either/or thinking. The term *psychoanalysis* refers to a great many things: a diverse collection of therapeutic techniques, a wide range of theories and models, a rich collection of diagnostic constructs, and a certain attitude toward mental life that is not wedded to any specific technique or model (Gabbard 2000; Schafer 1983). There are aspects of psychoanalysis that are best understood as hermeneutic. Certainly analytic listening, or the process by which we discern disavowed aspects of mental life in patients’ manifest communications, is largely hermeneutic. But other aspects of psychoanalysis fall under the categories of empirical questions and causal propositions. The theories that implicitly or explicitly guide analytic listening contain many causal propositions. We can employ hermeneutic methods where they facilitate understanding, and empirical methods where empirical research does.

A postmodern version of the argument that empirical research is irrelevant comes from analysts who regard the psychoanalytic encounter as an intersubjective construction in which analyst and analysand co-construct a unique reality. Some might go so far as to suggest that “character structure” is no longer a helpful concept and that there is therefore little point in assessing it. The intersubjective perspective has provided a useful corrective to the authoritarianism that has sometimes characterized analytic writings, and to the antiquated notion of analyst as blank screen. But the argument that we can disregard character structure seems indefensible at its core. However much we may value the uniqueness of our patients, we do not approach them in a vacuum. If we have expertise in treating human suffering, it is precisely because we can

recognize commonalities across patients and understand their suffering in the broader context of familiar patterns and syndromes.

Beyond these conceptual objections to research are many practical impediments to actually doing analytically meaningful research. The temperamental affinities of analysts and researchers differ considerably. Few who find fulfillment in the emotional crosscurrents and eddies of the psychoanalytic encounter also have the temperament of the dispassionate scientist. Becoming a psychoanalyst is a long and demanding process, and so is becoming a skilled researcher with an in-depth understanding of psychometrics and multivariate methods. Few attain mastery in both disciplines. Even where there is interest in empirical research, the practical necessity of making a living prevents many analysts from devoting significant time to it. One cannot become a sophisticated researcher in one's spare time, any more than one can become a sophisticated analyst. Finally, the professional culture of psychoanalysis has only recently begun to reward research. Those seeking professional advancement have therefore been motivated to pursue their ambitions in other ways. There are no obvious solutions to any of these problems, but perhaps there is some value in openly acknowledging them.¹

432

WHAT'S BEEN WRONG WITH EMPIRICAL RESEARCH?

Reproducible findings are the hallmark of empirical science. In their striving for reproducible data, researchers have sought to minimize reliance on clinical judgment and inference. This is the logic behind DSM-IV, which classifies personality pathology by counting manifest symptoms instead of considering dynamics and character organization. For example, the tendency to rely on projection as a defense is not included among the diagnostic criteria for paranoid personality disorder, in part because the framers of the manual assumed clinicians could not reliably assess defenses. The same logic underlies assessment instruments like the Structured Clinical Interview for DSM-IV (SCID;

¹Of course there are psychoanalytic investigators who have pursued empirical research despite these obstacles, including Enrico Jones, Lester Luborsky, Howard Shevrin, Robert Wallerstein, Peter Fonagy, Robert Emde, Daniel Stern, Sidney Blatt, Marianne Leuzinger-Bohleber, Horst Kächele, and others. For a review of psychoanalytic outcome studies see Fonagy et al. (2001).

First et al. 1997), which has become virtually de rigueur in psychiatric research. The interview renders irrelevant the skills of a dynamically sophisticated interviewer, because it is less a *clinical* interview than an interviewer-administered self-report questionnaire. For example, the SCID-II attempts to assess Narcissistic Personality Disorder by asking direct questions like “Do you feel that your situation is so special that you require preferential treatment?”

Such approaches diverge too far from clinical practice and reflect a too narrow view of “reproducibility.” Sophisticated clinicians do not limit themselves to information that is overt and manifest, or base diagnostic formulations exclusively on patients’ answers to direct questions. Patients with personality pathology generally lack the insight to describe their own personality characteristics. Narcissistic patients, for example, may well describe themselves as caring people and wonderful friends. Sophisticated clinicians *infer* traits such as grandiosity and entitlement from patients’ accounts of their important relationships, and from their behavior toward the clinician in the consulting room.

By eliminating clinical judgment and inference, or trying to reduce it to the lowest common denominator, researchers exclude from study precisely what is of greatest psychological interest. The resulting assessment instruments (i.e., highly structured interviews, rating scales, self-report questionnaires) do not capture the dynamically important aspects of mental life. I believe the only “instrument” sensitive enough to do so is the empathically attuned and dynamically sophisticated clinician given free rein to practice his or her craft. I will describe a research program aimed at *harnessing* clinical judgment rather than eliminating it. One goal of this research is to get this exquisitely sensitive “instrument,” the psychoanalytic clinician, to provide data in a form systematic enough for research purposes.

A NEW LANGUAGE FOR PSYCHOANALYTIC RESEARCH

Rulers measure in inches and scales measure in pounds, but what metric do analysts share? Imagine three analysts reviewing the same case material. One might speak of conflict and compromise formations, another of projected and introjected self- and object representations, and the third, perhaps, of self defects and fragmentation. The three analysts may be observing the same thing, but be using different language

and metaphor systems to express it. Or they may be attending to different aspects of the clinical material, à la the parable of the elephant and the blind men. Or they may not be able to make the same clinical observations at all. If we want to know whether the analysts can make the same observations, we must ensure that they speak the same language and pay attention to the same range of clinical phenomena.

The Shedler-Westen Assessment Procedure-200 (SWAP-200) is an assessment instrument designed to accomplish this (Shedler and Westen 1998; Westen and Shedler 1999a,b). It provides dynamic clinicians a common vocabulary with which to express their observations and inferences about character and character pathology. 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 eight categories, from those that are most descriptive of the patient (assigned a value of 7 for research purposes) to those that are not at all descriptive (assigned a value of 0). Thus, the SWAP-200 is not a patient questionnaire, but rather a method for systematically recording *clinicians'* observations.

434

The SWAP-200 yields a score from 0 to 7 for each of 200 personality-descriptive variables, but it is not the equivalent of a set of rating scales. The instrument is based on the Q-sort method, which requires clinicians to place a predetermined number of statements in each category. In psychometric terms, clinicians must arrange the statements into a “fixed distribution.”² The use of a fixed distribution has important psychometric advantages (Block 1976) and eliminates much of the measurement error or “noise” inherent in standard rating scales.³ Many psychoanalysts who attempt empirical research obtain disappointing

²The distribution for the SWAP-200 resembles the right half of a normal distribution or “bell-shaped curve.” Thus, 100 items are placed in the “0” or not descriptive category, and progressively fewer in higher categories. Only eight items are placed in the “7” or most descriptive category.

³One way it does so is by ensuring that clinicians 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 will tend to use extreme values (e.g., values of 0 and 7 on a 0–7 scale), while others of more moderate temperament will tend toward middle values (e.g., 4 and 5). Thus, the ratings reflect not just the psychological 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 1961).

results, not necessarily because their hypotheses are wrong, but because they do not understand psychometrics and their data contain too much measurement error to test *any* hypothesis (for other examples of the Q-sort method in psychoanalytically relevant research, see, e.g., Jones and Ablon 1998; Jones and Windholz 1990; Shedler and Block 1990).

The initial SWAP item set was drawn from a wide range of sources including contemporary psychoanalytic writings on character and character pathology (Shedler and Westen 1998; Westen and Shedler 1999a). The item set was revised and refined over a seven-year period through an iterative process. Approximately two hundred psychoanalytically trained clinicians used earlier versions of the instrument to describe a wide range of patients. Each time, we asked the clinicians one crucial question: “Were you able to describe the things you consider psychologically important about your patient?” We added, rewrote, and revised items based on this feedback, and then asked new clinicians to describe new patients. We repeated this process over many iterations until most clinicians could answer “yes” most of the time.⁴

435

The SWAP-200 items contain no jargon. We were determined that the statements not be couched in specialized language or metaphor systems meaningful only to adherents of particular theoretical traditions; in this respect, the approach is comparable to Schafer’s “action language” (1976). At the same time, we were determined not to exclude constructs that significant groups of analytic practitioners consider meaningful. Reconciling these goals was challenging. Many terms in analytic parlance tend to be used loosely, or used differently by different theorists (consider, for example, “projective identification” or “selfobject”). In these cases our practice was to articulate the multiple meanings of the terms, then write multiple items to cover the multiple meanings. I recommend to any analyst the exercise of translating familiar psychoanalytic constructs into jargon-free English. It is a humbling exercise that quickly exposes the limits of our understanding.

Because the SWAP-200 is jargon-free, it may have the potential to serve as a universal language for describing personality pathology.

⁴The empirical studies described in this article were conducted with Drew Westen. When I use the pronoun “we,” I am referring to our collaborative work.

Our studies demonstrate that experienced psychotherapists of diverse theoretical orientations can understand the items and apply them to their patients—whether or not they fully appreciate their dynamic significance or would otherwise have attended to the same aspects of the clinical material. In one study, a nationwide sample of 797 experienced psychologists and psychiatrists of diverse theoretical orientations, who had an average of eighteen years practice experience post-training, used the SWAP-200 to describe patients with personality pathology (Westen and Shedler 1999a). These experienced psychotherapists provided similar SWAP-200 descriptions regardless of their theoretical commitments, and fully 72.7 percent agreed with the statement “I was able to express most of the things I consider important about this patient” (the highest rating category).

AN ILLUSTRATION: BORDERLINE PERSONALITY ORGANIZATION

436

At first blush, it may seem that a finite set of 200 statements cannot begin to capture the richness and complexity of psychoanalytic formulations. However, the SWAP-200 statements can be combined in virtually infinite patterns to express subtle and sophisticated ideas. The musically inclined reader might consider that all of Western music can be notated using combinations of only twelve tones. The mathematically inclined might note that the number of possible orderings of the SWAP-200 statements is 200 factorial, which is an inexpressibly large number. Consider, for example, the SWAP-200 items reproduced below. The first three items, taken in *combination*, convey something of the defensive splitting seen in patients with borderline character organization:

- 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 next three items, taken in combination, capture at least one of the meanings of the term *projective identification*:

- 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 he or she 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 elicit extreme reactions or stir up strong feelings in others (e.g., others may become uncharacteristically enraged, anxious, depressed, etc. when interacting with him/her).

The next group of items helps flesh out a picture of a certain kind of borderline patient, addressing issues of affect regulation, identity diffusion, object relations, self experience, and ego boundaries (in a way consistent with the descriptions of Kernberg [1984; Kernberg et al. 1989]):

- 191 Emotions tend to change rapidly and unpredictably.
- 12 Emotions tend to spiral out of control, leading to extremes of anxiety, sadness, rage, excitement, etc.
- 157 Tends to become irrational when strong emotions are stirred up; may show a noticeable decline from customary level of functioning.
- 16 Tends to express intense and inappropriate anger, out of proportion to the situation at hand.
- 15 Lacks a stable image of who s/he is or would like to become (e.g., attitudes, values, goals, and feelings about self may be unstable and changing).
- 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.
- 98 Tends to fear s/he will be rejected or abandoned by those who are emotionally significant.
- 167 Is simultaneously needy of, and rejecting toward, others (e.g., craves intimacy and caring, but tends to reject it when offered).
- 153 Interpersonal relationships tend to be unstable, chaotic, and rapidly changing.
- 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; often misunderstands, misinterprets, or is confused by others' actions and reactions.

- 176 Tends to confuse own thoughts, feelings, or personality traits with those of others (e.g., may use the same words to describe the self and another person, believe the two share identical thoughts and feelings, treat the other as an “extension” of him/herself, etc.).

The next group of items includes descriptors that might apply to a more disturbed type of borderline patient, perhaps one likely to be seen in an inpatient setting (see Gunderson 1984).

- 134 Tends to act impulsively, without regard for consequences.
109 Tends to engage in self-mutilating behavior (e.g., self-cutting, self-burning, etc.).
142 Tends to make repeated suicidal threats or gestures, either as a "cry for help" or as an effort to manipulate others.
122 Living arrangements tend to be chaotic or unstable (e.g., living arrangements are temporary, transitional, or ill-defined; may have no telephone or permanent address).
188 Work life tends to be chaotic or unstable (e.g., working arrangements seem always temporary, transitional, or ill-defined).
44 Perception of reality can become *grossly* impaired under stress (e.g., may become delusional).

438

The items reproduced above are illustrative only, and are not intended to describe “the” borderline patient or even any particular borderline patient. Other SWAP-200 items could add many flavors and nuances to the description, including specific defensive processes, areas of conflict, transference propensities, types of self disturbance, object-relational configurations, and so on, to create a highly individualized portrait. The items are intended only to illustrate that dynamic processes can be described without recourse to jargon and (I hope) to persuade the reader that the SWAP-200 is not academic research “business as usual,” but rather a genuine effort to operationalize psychoanalytically meaningful constructs.

ARE PSYCHOANALYTIC OBSERVATIONS RELIABLE?

Reproducible findings are crucial to any discipline that would claim scientific respectability. As I noted earlier, however, the term *psychoanalysis* encompasses many things, and there is therefore much room for debate about just what kind of “findings” are relevant and when. I will restrict my attention to the most basic kind of findings: clinical observations.

Consider an analogy. When radiologists read X rays, they are making clinical observations. We would have little faith in radiology if independent radiologists could not agree with one another, at least about routine observations. Similarly, when we say that a patient has a narcissistic character organization, or a borderline character organization, or relies on splitting as a defense, or has developed an erotic transference, or has malevolent object representations, we too are making clinical observations. These are the routine observations of analytic practice. These observations are either reliable or not. If not, then our critics have every reason to regard our pronouncements with skepticism.

Freud (1916–1917) explicitly recognized the importance of this kind of reliability. In discussing the technique of dream interpretation, he anticipated the criticism that “so much room is left to the interpreter’s arbitrary decisions as to be incompatible with objective certainty” (p. 228). Freud’s response to the hypothetical critic was that we can establish the soundness of the psychoanalytic method “if we *compare the results* produced by correctly trained analysts” (p. 232; emphasis added). Freud made an analogy with the deciphering of ancient cuneiform, which had been widely regarded as a hoax until a “decisive experiment” (p. 232) showed that independent scholars could arrive at similar translations. A contemporary researcher would eschew the phrase “objective certainty,” but Freud’s logic is exactly the logic of contemporary empirical research; Freud was describing interrater reliability.

The SWAP-200 provides the technology necessary to perform the kind of “decisive experiment” Freud advocated, with respect to analytic observations about character organization. Because it translates analytic formulations into quantitative form, it allows us to compare the findings of independent analysts and measure interrater reliability. A SWAP-200 description of a patient consists of one column by 200 rows of data, each row containing the score (0 to 7) for the corresponding SWAP-200 item. Two SWAP-200 descriptions yield two columns of data. We can correlate the two columns to measure the overall agreement or similarity between the two descriptions.⁵

What is the result of our “decisive experiment”? Our research demonstrates that psychoanalytic clinicians can make highly reliable

⁵The correlation reflects the pattern or configuration across all 200 items, not the reliability of any single item.

observations and inferences about character organization. In our studies, analytically oriented clinicians, working independently of one another, either conducted or observed (by videotape) a series of extended intake interviews (the equivalent of a three-hour intake process, modeled after clinical practice), and then described the patient using the SWAP-200. Reliability coefficients ranged from $r = .75$ to $r = .81$ (Shedler and Westen 1998).⁶ We observed similarly high correlations between clinicians who conducted independent diagnostic interviews and the patients' own therapists, who developed their clinical formulations over many months of clinical work. These reliability coefficients are higher than those typically reported by psychiatric researchers using highly structured interviews that just "stick to the facts" (i.e., DSM criteria).

In short, psychoanalytic clinicians can make reliable observations and inferences about personality organization, provided they have a suitable technology for harnessing their judgments. The widely held view among researchers, that psychoanalytic constructs cannot be assessed reliably, is mistaken.

440

OTHER CLINICAL AND EMPIRICAL APPLICATIONS

Case Formulation and Clinical Supervision

One application of the SWAP-200 is in clinical training and supervision. To describe a patient using the SWAP-200, a clinician must attend to multiple facets of the clinical material and consider the patient from multiple vantage points (e.g., those of drive, ego, object, and self; see Pine 1990). This is so because the SWAP-200 items reflect various theoretical perspectives, and the clinician must render a decision about each item, if only to judge it inapplicable.

A supervisee told me her patient had "fragmented" during a session. When I asked what she meant by that, she was unable to articulate her thoughts. She had been accustomed to using the term with a previous supervisor, and the two had shared the unexamined (and apparently erroneous) assumption that they understood the term the same way. I asked specific questions to help her clarify her experience of the patient. Were the patient's attitudes and beliefs unstable and changing?

⁶These coefficients represent the reliability of the composite SWAP-200 description, after the scores provided by the two clinicians are averaged together (Spearman-Brown formula).

Did she lose a sense of the continuity of her experiences over time? Did she lose a sense of bodily integrity or coherence? Did she confuse her thoughts and feelings with those of others? As it turned out, my supervisee meant none of these things. After much discussion we both realized that the patient had simply become *anxious*, and this had prompted my supervisee to offer reassurances. Once we clarified that the patient was anxious and not “fragmenting,” my supervisee stopped offering reassurances and instead began encouraging her to articulate her fears. This yielded new material, and the treatment took a more productive turn. Subsequently, I asked my supervisee to describe her patient using the SWAP-200. As she considered each item in turn, she found herself engaged in the kind of reasoning I had encouraged in supervision. The exercise fostered a process of self-supervision that led to clearer thinking and more precise formulations.

Assessing Structural Change

Some findings from psychotherapy outcome studies make little sense. Oft-cited findings that have deviled psychotherapy researchers are that (a) all forms of therapy achieve equivalent results; (b) that all therapists, regardless of training or experience, achieve equivalent results; and (c) that psychotherapy and antidepressant medication achieve equivalent results (see, e.g., Elkin et al. 1989; Lambert and Bergin 1994; Luborsky et al. 1993; VandenBos 1996). Of course no psychoanalytic clinician really believes these things (I doubt any working clinician believes them). Still, we have not empirically demonstrated that psychoanalytic therapies accomplish anything that cannot be accomplished by briefer and more superficial therapies. Why haven't we?

One problem lies in the outcome measures used in the studies.⁷ Virtually all published psychotherapy outcome studies (and psychotropic medication trials) define favorable outcome in terms of symptom remission. The outcome measures do not attempt to assess structural change or the positive presence of psychological strengths that may be developed through treatment. It appears that many forms of therapy can achieve remission of manifest symptoms, at least in the short run. I believe no other form of therapy achieves the kind of

⁷A second problem is that psychoanalysts have rarely published empirical outcome studies. Most published studies have been conducted by cognitive-behavioral and other nonanalytic investigators, who understandably wish to demonstrate the advantages of their own approaches.

lasting structural changes we see when psychoanalytic treatments go well. I am referring to changes in characteristic compromise formations, less rigid use of defenses, a shift toward more mature defenses, changes in object relations and object representations (e.g., more integrated or more benign representations), more coherent experience of self, and so on. But at present this is just the opinion of one investigator drawn to psychoanalysis by his own training, clinical experience, and personal dynamics. With the right outcome measure we could empirically evaluate the benefits of psychoanalytic therapies relative to other therapies.

The SWAP-200 provides such an outcome measure. Because it assesses personality structure or organization, it can tap the psychological phenomena that actually change in psychoanalytic psychotherapy. In addition, the instrument includes approximately thirty items that reflect the positive presence of psychological strengths. Consider the following partial listing of “healthy” SWAP-200 items:

442

- 2 Is able to use his/her talents, abilities, and energy effectively and productively.
- 32 Is capable of sustaining a meaningful love relationship characterized by genuine intimacy and caring.
- 37 Finds meaning in belonging and contributing to a larger community (e.g., organization, church, neighborhood, etc.).
- 55 Is able to find meaning and fulfillment in guiding, mentoring, or nurturing others.
- 59 Is empathic; is sensitive and responsive to other peoples' needs and feelings.
- 68 Appreciates and responds to humor.
- 82 Is capable of hearing information that is emotionally threatening (i.e., that challenges cherished beliefs, perceptions, and self-perceptions) and can use and benefit from it.
- 89 Appears to have come to terms with painful experiences from the past; has found meaning in, and grown from, such experiences.
- 111 Has the capacity to recognize alternative viewpoints, even in matters that stir up strong feelings.
- 196 Is able to find meaning and satisfaction in the pursuit of long-term goals and ambitions.
- 200 Is able to form close and lasting friendships characterized by mutual support and sharing of experience.

It is doubtful that antidepressant medication or twelve sessions of cognitive-behavioral therapy would significantly change the ranking of some of these items in a patient's SWAP-200 description. I think it likely that psychoanalysis would. We could investigate this, by designing outcome studies in which patients participate in videotaped clinical interviews, and panels of independent clinicians (not the treating therapists) assess the patient's character organization pre- and post-treatment using the SWAP-200. Such an approach might enable us, at last, to empirically demonstrate the unique benefits of psychoanalytic treatment. If health insurers did not take note of the findings, no doubt many patients and prospective patients would.

TOWARD DSM-V

It is no secret that the personality disorder taxonomy provided by DSM-IV has been of limited relevance to psychoanalytic clinicians. McWilliams (1999) has stated the problem plainly:

As succeeding editions of the Diagnostic and Statistical Manual of Mental Disorders (DSM) of the American Psychiatric Association have become increasingly objective, descriptive, and putatively atheoretical, they have inevitably minimized the subjective and inferential aspects of diagnosis on which most clinicians actually depend. Operating more or less invisibly alongside the empirically derived categories of the DSM is another compendium of wisdom, passed down orally and in practice-oriented journals: clinical knowledge, complexly determined inferences, and consistent impressions made on the harnessed subjectivities of therapists. In any individual case, these data coexist somewhat uneasily with whatever formal diagnostic label the patient has been given [p. 1].

443

In short, the DSM taxonomy does not reflect the way working clinicians think. Nor does it rest on a firm empirical foundation. In fact, its categories and criteria are not empirically derived. Rather, they are the product of committee decision processes, and are often at odds with empirical findings. There is excessively high comorbidity between the diagnostic categories, such that patients who meet the criteria for one personality disorder often meet the criteria for four to six. The test-retest reliability of Axis II diagnoses is poor after intervals of six weeks, despite efforts to make the diagnostic criteria "objective" (for a review and critique, see Westen and Shedler 1999a). Moreover,

approximately 60 percent of patients in treatment for personality pathology cannot be diagnosed on Axis II at all (Westen and Arkowitz-Westen 1998).

These problems and others led us to pursue a research program aimed at developing an alternative personality disorder taxonomy—one that is clinically near, dynamically sensible, and empirically grounded. The SWAP-200 has been the primary vehicle for this research. Because it quantifies dynamic formulations, it allows us to study personality pathology in ways not previously possible. I believe it can provide a bridge between descriptive psychiatry and dynamic case formulation.

In one study, a nationwide sample of 496 psychologists and psychiatrists (with diverse theoretical orientations) used the SWAP-200 to describe a current patient with a personality disorder. We employed a statistical technique known as Q-analysis to examine the resulting data to search for naturally occurring clusters or groupings of patients—i.e., patients who share important psychological features that differentiate them from other patients. (Q-analysis is computationally equivalent to the more familiar technique of factor analysis; the difference is that factor analysis searches for groups of similar variables [columns in a data table], whereas Q-analysis searches for groups of similar people [cases or rows in a data table]). The technique has been used by biologists to aid in the classification of species, and by academic psychologists to study normal personality (Block 1971). In essence, we set aside all theoretical preconceptions and simply let the data speak.

We have published a detailed description of our methods and findings elsewhere (Westen and Shedler 1999b), but I will summarize some relevant results here. The Q-analysis identified eleven diagnostic groups (natural taxonomic categories) based on similar dynamics and character organization, not simply on manifest symptoms. We created a composite personality description or “prototype” for each diagnostic group by averaging the SWAP-200 descriptions of patients in the group. The SWAP-200 items with the highest scores in the prototype reveal the salient and defining features of the disorder.

Our empirically derived diagnostic groups bore a family resemblance to Axis II categories (many of which in turn have their origins in psychoanalytic tradition), in that there were groups of patients whose personality organizations were clearly hysterical, obsessional, paranoid, narcissistic, and so on. But our composite descriptions were clinically richer than the DSM criterion sets, and more closely reflected

the “compendium of wisdom” passed down from generation to generation of analytic clinicians (see, e.g., Shapiro 1965; MacKinnon and Michels 1971; McWilliams 1994; see also Akhtar 1992).

In particular, the composite personality descriptions included items addressing dynamic and intrapsychic factors not acknowledged by the DSM. The findings made clear that psychodynamic factors are central to an understanding of personality pathology, and are observed even by clinicians not committed to a psychoanalytic perspective. For example, the composite SWAP-200 description of patients with paranoid personality organization included not only manifest suspiciousness, but also aggression, feelings of victimization, and the defenses of externalization and projection:

- 87 Is quick to assume that others wish to harm or take advantage of him/her; tends to perceive malevolent intentions in others' words and actions.
- 185 *Is prone to intense and inappropriate anger, out of proportion to the situation at hand.*
- 127 *Tends to feel misunderstood, mistreated, or victimized.*
- 14 *Tends to blame others for own failures or shortcomings; tends to believe his/her problems are caused by external factors.*
- 116 *Tends to see own unacceptable feelings or impulses in others instead of in him/herself.⁸*

445

The composite SWAP-200 description of patients with histrionic personality organization emphasized not only manifest emotionality, sexual attention seeking, suggestibility, etc., but also underlying neediness and dependency. These latter features are consistent with the observations of many analytic writers but absent from the DSM criterion set:

- 74 Expresses emotion in exaggerated and theatrical ways.
- 97 Tends to use his/her physical attractiveness to an excessive degree to gain attention or notice.
- 34 Tends to be overly sexually seductive or provocative, whether consciously or unconsciously (e.g., may be inappropriately flirtatious, preoccupied with sexual conquest, prone to “lead people on,” etc.).
- 18 Tends to choose sexual or romantic partners who seem inappropriate in terms of age, status (e.g., social, economic, intellectual), etc.

⁸The SWAP-200 items reproduced here are illustrative only, and are not full descriptions of the personality syndromes. For complete descriptions, see Westen and Shedler (1999b).

- 46 Tends to be suggestible or easily influenced.
- 22 Tends to develop somatic symptoms in response to stress or conflict (e.g., headache, backache, abdominal pain, asthma, etc.).
- 77 *Tends to be overly needy or dependent; requires excessive reassurance or approval.*
- 98 *Tends to fear he/she will be rejected or abandoned by those who are emotionally significant.*

The composite SWAP-200 description of patients with narcissistic personality organization emphasized not only manifest grandiosity and inflated self-worth, but also underlying emptiness, a false self, perfectionistic superego introjects, and the need for others to serve as selfobjects:

- 49 Has fantasies of unlimited success, power, beauty, talent, brilliance, etc.
- 190 Appears to feel privileged and entitled; expects preferential treatment.
- 4 Has an exaggerated sense of self-importance.
- 133 Tends to be arrogant, haughty, or dismissive.
- 143 Tends to believe s/he can only be appreciated by, or should only associate with, people who are high-status, superior, or otherwise “special.”
- 50 *Tends to feel life has no meaning.*
- 38 *Tends to feel s/he is not his/her true self with others; tends to feel false or fraudulent.*
- 174 *Expects self to be “perfect” (e.g., in appearance, achievements, performance, etc.).*
- 53 *Seems to treat others primarily as an audience to witness own importance, brilliance, beauty, etc.*
- 128 *Fantasizes about finding ideal, perfect love.*

A crucial finding was the prevalence of a form of personality organization that is not acknowledged by DSM at all—depressive personality. Indeed, the largest percentage of patients in our sample (over 20 percent) belonged to this diagnostic group. No doubt our empirical “discovery” of depressive personality comes as no surprise to psychoanalytic clinicians, since the psychoanalytic literature is filled with descriptions of depressive personality dynamics. However, such empirical data may have implications for future revisions of DSM, and may provide a corrective to the view of depression, promoted by pharma-

ceutical companies, health insurers, and biological psychiatry, as exclusively a medical illness or “brain disease.”

Along with many symptom-oriented items, the SWAP-200 description of depressive personality contained items of considerable relevance to psychoanalytic theory. They included items addressing active masochism and the centrality of anger (consider the meaning of items 16, 25, and 78, taken in combination):

- 189 Tends to feel unhappy, depressed, or despondent.
- 56 Appears to find little or no pleasure, satisfaction, or enjoyment in life's activities.
- 1 Tends to blame self or feel responsible when bad things happen.
- 163 *Appears to want to “punish” self; creates situations that lead to unhappiness, or actively avoids opportunities for pleasure and gratification.*
- 16 *Tends to be angry or hostile (whether consciously or unconsciously).*
- 25 *Has difficulty acknowledging or expressing anger.*
- 78 *Tends to express aggression in passive and indirect ways (e.g., may make mistakes, procrastinate, forget, become sulky, etc.).*

447

These items are not taken from the description of any one patient, but from an empirically generated composite of many patients.⁹ Nor can the items be explained away as an artifact of the theoretical views of a “biased” sample of psychoanalysts, since they appeared also in SWAP-200 descriptions provided by cognitive-behavioral therapists and biologically oriented psychiatrists. In short, these characteristics are central to the dynamics of many depressive patients, as observable today as they were by Freud (1917) more than eighty years ago. Thus, the SWAP-200 provides a bridge between the merely descriptive constructs employed by empirical depression researchers and the dynamic and explanatory constructs sought by psychoanalytic clinicians.

I want to emphasize several differences between our approach and that of DSM. First, our diagnostic categories and “criteria” are empirically derived and therefore unarguably faithful to the data. To the extent allowed by contemporary data analysis methods, they reflect the categories that exist in nature. By contrast, the DSM categories and criteria

⁹These patients could plausibly have been described as having masochistic personality, but we reserved the term *masochistic* for another diagnostic group identified in our analysis, in which masochism was more salient and defining.

reflect the opinions of committees, which can be influenced by group dynamics, the personalities of committee members, the sociopolitical zeitgeist, and other such factors. Thus, our inclusion of a depressive personality category is based not on opinion or aesthetic preference, but on evidence. The SWAP-200 therefore provides the technology to establish a personality disorder taxonomy on a stronger empirical foundation than DSM. Second, the method allows us to conceptualize personality disorders as continua, not categories. For example, clinicians can describe borderline pathology on a continuum ranging from mild through moderate to severe, rather than classifying borderline personality disorder as present/absent. We believe this more accurately reflects clinical thinking. Third, the SWAP-200 incorporates intrapsychic and dynamic factors such as motives, fantasies, object representations, conflict, and defense, whereas DSM emphasizes manifest symptoms. One can think of DSM diagnostic criteria as “building blocks” that allow us to model the external facade of a personality disorder. The SWAP-200 allows us to model the interior as well.¹⁰

448

The study described above is but a preliminary step in a larger research project aimed at developing an alternative to the DSM personality disorder taxonomy. We have received funding from the National Institute of Mental Health to conduct further research using larger and more diverse patient samples. We have also developed an interactive, web-based version of the SWAP-200, which automates the data-recording procedure and makes the instrument accessible to clinicians worldwide.¹¹ The research may help put future revisions of DSM on a stronger empirical foundation, and also make it more clinically and dynamically relevant. If we are successful, mainstream psychiatric diagnosis will once again reflect dynamic thinking.

A CASE ILLUSTRATION: MR. N.

I have emphasized that the SWAP-200 is designed to provide a bridge between the worlds of psychoanalysis and empirical research. The SWAP-200 has strong empirical credentials (Shedler and Westen 1998; Westen and Shedler 1999a,b). In writing this article, however, I have struggled with the feeling that psychoanalytic readers may

¹⁰I thank Meriamne Singer for suggesting this analogy.

¹¹To preview the instrument on the web, visit www.psychsystems.net/guest.cfm

remain unconvinced that the SWAP-200 is truly useful for case formulation. Can the SWAP-200 “bring a patient to life” and contribute to a meaningful dynamic understanding? I will let the reader judge.

Mr. N., a forty-eight-year-old white male, was selected from the approximately five hundred patients in our study (Westen and Shedler 1999b). He had been seen for nine psychotherapy sessions at the time his analyst described him using the SWAP-200. He is college-educated, and his analyst regards him as high-functioning (based on the Global Assessment of Functioning score the analyst provided). On our questionnaires, the analyst indicated that Mr. N. had suffered no noteworthy childhood traumas, although he rated Mr. N.’s relationship with his father as very poor. Mr. N.’s genetic history is unremarkable.

The following is a list of the SWAP-200 items that Mr. N.’s analyst placed in the three highest, or most descriptive, categories (i.e., items receiving scores of 5, 6, or 7). The items are reproduced essentially verbatim, with only minor grammatical changes to aid the flow of the text.

Mr. N. has an exaggerated sense of self-importance; feels privileged and entitled; believes he can only be appreciated by, or should only associate with, people who are high-status, superior, or otherwise “special”; fantasizes about unlimited success, power, beauty, talent, brilliance, etc.; treats others primarily as an audience to witness his own importance, brilliance, beauty, etc.; seeks to be the center of attention; tends to be arrogant, haughty, and dismissive; and feels that an important other has a special, seemingly magical ability to understand his innermost thoughts and feelings (e.g., he may imagine that rapport is so perfect that ordinary communication is superfluous).

He tends to be angry or hostile (whether consciously or unconsciously), tends to be controlling, and tends to be conflicted about authority (e.g., he may feel he must submit, rebel, win over, defeat). He tends to express aggression in passive and indirect ways (e.g., by making mistakes, procrastinating, forgetting, becoming sulky). He is prone to think in abstract and intellectualized terms, even in matters of personal import. He repeatedly convinces others of his commitment to change, only to revert to his previous maladaptive behavior (i.e., he convinces people that “this time is really different”).

Mr. N. fantasizes about finding ideal, perfect love; uses his physical attractiveness to an excessive degree to gain attention and notice; tends to be overly sexually seductive or provocative, whether consciously or unconsciously (e.g., he may be inappropriately flirtatious, preoccupied with

sexual conquest, or prone to “lead people on”); tends to be hostile toward members of the opposite sex, whether consciously or unconsciously; and appears afraid of commitment to a long-term love relationship.

Along with his pathology, Mr. N. has considerable psychological strengths. He is energetic and outgoing; tends to be liked by others; has an active and satisfying sex life; is articulate; appreciates and responds to humor; appears comfortable and at ease in social situations; is creative and able to approach problems in novel ways; can assert himself effectively and appropriately when necessary; and appears to have come to terms with painful experiences from the past, having found meaning in, and grown from, such experiences.

Case Formulation

From this configuration of psychological characteristics, we can draw some inferences. Mr. N. is a high-functioning narcissistic character. He can be charming and likeable, and he uses his charm to win admiration and affection. At the same time, he is self-centered, feels entitled, and values others primarily to the extent that they bolster his grandiose but fragile view of himself (e.g., by offering admiration or witnessing his magnificence). His relationships begin with promise, only to sour with time. These dynamics find particular expression in Mr. N.’s relations with women. I suspect he is a womanizer who leaves victims in his wake, because he is charming and leads women on, but is unable to sustain a meaningful relationship characterized by mutual empathy, caring, and sharing. He cannot do so, in part because at the core he seeks someone whose role is to help regulate his self-esteem by understanding him perfectly, admiring his perfection, and being perfect herself. He is angry and subtly devaluing toward women, who fail to fulfill these expectations. As he nears fifty, however, his fantasies about ideal, perfect love are increasingly difficult to sustain. I suspect he is confused and pained by his repeated failed relationships, and this may well be what has brought him to treatment.

If these inferences are correct, we can expect Mr. N. to express these personality dynamics in the transference. He might seek out a therapist he can regard as special and superior, like himself, who will share in his perfection and understand him perfectly (a “twinship” transference); and/or he may devalue the therapist, who must ultimately frustrate and disappoint. He may attempt to seduce or coerce the analyst into providing narcissistic supplies, or become enraged when he

perceives the analyst as withholding them. The analyst's job will be to help Mr. N. recognize these patterns and the functions they serve, instead of simply enacting them with a new object. A classical analyst might emphasize conflict (e.g., between dependency and rage) and its manifestations in the transference and other important relationships. An analyst with an object-relational perspective might emphasize Mr. N.'s projection of both idealized and devalued self-representations, and the therapeutic value of containing and helping "metabolize" these projections. A self psychologist might emphasize Mr. N.'s difficulty in regulating self-esteem and his use of the analyst as a selfobject to provide this regulatory function.

Mr. N.'s tendency to intellectualize will doubtless present difficulties in the treatment, because he may treat interpretations as abstract theories to ponder, without the personal relevance and affective charge that lead to change. The analyst will need to address the intellectualization when it arises, lest the treatment become an academic exercise. The relative absence of anxiety and depressive affect lead me to wonder whether Mr. N. will sustain his commitment to treatment over time. However, the fact that Mr. N. has found meaning in past painful experiences and grown from them, along with his considerable psychological resources, increases the likelihood that he will use treatment effectively.

451

Some Comments

I leave it to the reader to judge whether this description of Mr. N. is consistent with the spirit of dynamic case formulation. I do, however, want to emphasize two points. First, with the SWAP-200, descriptive psychiatric diagnosis and dynamic case formulation become part of the same process. We can use SWAP-200 data to objectively assess the correlation or "fit" between an individual patient and a diagnostic prototype. For example, Mr. N.'s SWAP-200 description correlates highly with the narcissistic personality disorder prototype, and moderately with the obsessional and histrionic prototypes. Thus, in DSM parlance his diagnosis would be "narcissistic personality disorder with obsessional and histrionic features" (Westen and Shedler 1999b). At the same time, the SWAP-200 provides a narrative case description from which we can derive dynamic formulations relevant to treatment. Descriptive psychiatry and dynamic case formulation need not, therefore, be unrelated activities.

Second, the standard vocabulary of the SWAP-200 ensures that different analysts will describe a patient in much the same way, once they become familiar with the instrument. Had another analyst described Mr. N. using the SWAP-200, the narrative description would have been much the same, because every word of the description was taken directly from the SWAP-200 item set. It is this quality, I believe, that makes the SWAP-200 a unique tool for psychoanalytic research. With the problem of reliability (reproducibility of findings) solved, many aspects of psychoanalysis become open to empirical inquiry. Such inquiry can only strengthen our discipline.

DISCUSSION

452

There has long been a rift between psychoanalytic and empirical approaches to the study of personality pathology. Empirical researchers have too often disregarded intrapsychic and dynamic factors, either because they considered them unimportant or because they assumed dynamic factors could not be assessed reliably. In general, researchers have attempted to eliminate clinical judgment from assessment procedures, or else reduce clinical judgment to the lowest common denominator.

Because empirical researchers have so often disregarded dynamic factors, many analysts have come to believe that it is impossible to operationalize psychoanalytic constructs, or even that efforts to do so can serve only to distort or trivialize them (e.g., Green 1996, 2000). Thus, the dichotomy between empirical and psychoanalytic approaches has been perpetuated by researchers and analysts alike.

I believe the dichotomy is, in many instances, false. It derives from a narrow view of "reliability" or "reproducibility," based on the mistaken assumption that reliability can be achieved only by eliminating clinical subjectivity from psychological assessment. The assumption is mistaken for two reasons. First, it is empirically false. As DSM diagnostic criteria have become increasingly "objective," personality disorder diagnoses have not, as a result, become more reliable (see, e.g., Perry 1992; Zimmerman 1994). Second, elimination of human inference is not a requirement of science. Contemporary philosophers of science recognize that the observed, the observer, and the context of observation are thoroughly interdependent.

Earlier I used the analogy of radiologists reading X rays to emphasize the importance of reproducible findings. Embedded in the analogy is a second truth that too often has been lost on psychological and psychiatric researchers. This second truth is that “reproducible” need not mean reproducible by just anyone. We do not expect untrained lay people to read X rays, because the task, by its nature, requires expert judgment. The same holds true in psychology and psychiatry. We should not expect individuals who lack adequate training, experience, and sophistication to make correct dynamic observations and inferences. Science can encompass “complexly determined inferences and diagnostic impressions made on the harnessed subjectivities” of expert clinicians (McWilliams 1999, p. 1). Efforts to eliminate clinical judgment and inferences have not made psychological and psychiatric research more “scientific,” just more shallow.

Some analysts may feel that something vital to psychoanalysis will be lost by admitting the use of any standardized instrument. For example, certain gifted clinicians can describe a patient in language so moving and evocative that they create a sense of empathic resonance with the patient. Of course, the SWAP-200 cannot reproduce the poetry of such a case description. Nor can it capture all the personal, private, and idiosyncratic details of an individual patient’s life. It can describe areas of conflict, but not the deeply personal ways in which an individual patient experiences the conflict. It does not address developmental dynamics or genetic origins.

None of this should detract from what the SWAP-200 *can* do. It ensures that an analyst will attend to a very broad spectrum of relevant clinical material and consider a patient from multiple vantage points. It ensures that the analyst will express his or her observations in clear and unambiguous language. It provides an objective means of gauging the degree to which a patient’s personality dynamics match those of recognized personality syndromes. It provides an objective means of assessing change over the course of treatment. Not least, it has generated strong empirical data that will make it hard for psychiatric researchers to ignore psychodynamic factors in the future.

These are not necessarily bad trade-offs, but we do not have to make *any* trade-offs. Nothing in the approach precludes traditional methods of analytic inquiry, such as narrative case study; it is simply another tool in the epistemic arsenal of psychoanalysis. Indeed, SWAP-200 data can supplement and enhance narrative case material. For

example, one could use SWAP-200 items to anchor a case presentation with clear and scientifically meaningful markers, while weaving narrative, historical, and interpretive information around those markers (see Jones and Windholz 1990).

Throughout I have tried to make the case that it is possible to do psychoanalytically relevant research, and that there is every reason to do it. Although some analysts may object to empirical research on philosophical grounds, my sense is that the chief impediment to psychoanalytic research is the difficulty of operationalizing psychoanalytic constructs without eviscerating them. The SWAP-200 attempts to solve this problem with respect to case description. It provides a “language” for psychoanalytic research 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 researchers. There remains a sizeable gulf between psychoanalysts and empirical researchers. Perhaps this new language will be one all parties can speak.

REFERENCES

454

- AKHTAR, S. (1992). *Broken Structures: Severe Personality Disorders and Their Treatment*. Northvale, NJ: Aronson.
- BLOCK, J. (1961). *The Q-sort Method in Personality and Psychiatric Research*. Palo Alto: Consulting Psychologists Press, 1978.
- (1971). *Lives through Time*. Berkeley: Bankcroft Books.
- BRENNER, C. (1982). *The Mind in Conflict*. New York: International Universities Press.
- CHAMBLESS, D.L. (1996). In defense of dissemination of empirically supported psychological interventions. *Clinical Psychology: Science and Practice* 3:230–235.
- EAGLE, M. (1987). *Recent Developments in Psychoanalysis*. Cambridge: Harvard University Press.
- ELKIN, I., SHEA, M.T., WATKINS, J., & IMBER, S. (1989). National Institute of Mental Health Treatment of Depression Collaborative Research Program: General effectiveness of treatments. *American Journal of Psychiatry* 46:971–982.
- FIRST, M., SPITZER, R., GIBBON, M., & WILLIAMS, J. (1997). *Structured Clinical Interview for DSM-IV Personality Disorders (SCID-II)*. Washington, DC: American Psychiatric Press.
- FONAGY, P., KACHELE, H., KRAUSE, R., JONES, E., & PERRON, R. (2000). *An Open Door Review of Outcome Studies in Psychoanalysis*. Electronic publishing: <http://www.ipa.org.uk/research/R-outcome.htm>.
- FREUD, S. (1916–1917). Introductory lectures on psychoanalysis. *Standard*

- Edition 15/16.
- (1917). Mourning and melancholia. *Standard Edition* 14:237–258.
- GABBARD, G. (2000). *Psychodynamic Psychiatry in Clinical Practice*. 3rd ed. Washington, DC: American Psychiatric Press.
- GEDO, J.E. (1980). Reflections on some current controversies in psychoanalysis. *Journal of the American Psychoanalytic Association* 28:363–383.
- GREEN, A. (1996). What kind of research for psychoanalysis? *International Psychoanalysis* 5:8–9.
- (2000). Response to Robert S. Wallerstein. In *Clinical and Observational Psychoanalytic Research: Roots of a Controversy*, ed. J. Sandler, A.-M. Sandler, & R. Davies. Madison, CT: International Universities Press, pp. 32–37.
- GUNDERSON, J.G. (1984). *Borderline Personality Disorder*. Washington, DC: American Psychiatric Press.
- JONES, E.E., & ABLON, J.S. (1998). How expert clinicians' prototypes of an ideal treatment correlate with outcome in psychodynamic and cognitive-behavioral therapy. *Psychotherapy Research* 8:71–83.
- & WINDHOLZ, M. (1990). The psychoanalytic case study: Toward a method for systematic inquiry. *Journal of the American Psychoanalytic Association* 38:985–1009.
- KERNBERG, O. (1984). *Severe Personality Disorders*. New Haven: Yale University Press.
- SELZER, M.A., KOENIGSBERG, H.W., CARR, A.C., & APPELBAUM, A.H. (1989). *Psychodynamic Psychotherapy of Borderline Patients*. New York: Basic Books.
- LAMBERT, M.J., & BERGIN, A.E. (1994). The effectiveness of psychotherapy. In *Handbook of Psychotherapy and Behavior Change*, ed. A. Bergin & S. Garfield. New York: Wiley.
- LUBORSKY, L., DIGUER, L., LUBORSKY, E., ET AL. (1993). The efficacy of dynamic psychotherapies: Is it true that “everyone has won and all must have prizes”? In *Psychodynamic Treatment Research*, ed. N.E. Miller et al. New York: Basic Books, pp. 497–516.
- MACKINNON, R., & MICHELS, R. (1971). *The Psychiatric Interview in Clinical Practice*. Philadelphia: Saunders.
- MCWILLIAMS, N. (1994). *Psychoanalytic Diagnosis: Understanding Personality Structure in the Clinical Process*. New York: Guilford Press.
- (1999). *Psychoanalytic Case Formulation*. New York: Guilford Press.
- PERRY, J.C. (1992). Problems and considerations in the valid assessment of personality disorders. *American Journal of Psychiatry* 149:1645–1653.
- PINE, F. (1990). *Drive, Ego, Object, and Self: A Synthesis for Clinical Work*. New York: Basic Books.
- (1998). *Diversity and Direction in Psychoanalytic Technique*. New

Haven: Yale University Press.

SCHAFFER, R. (1976). *A New Language for Psychoanalysis*. New Haven: Yale University Press.

——— (1983). *The Analytic Attitude*. New York: Basic Books.

SHAPIRO, D. (1965). *Neurotic Styles*. New York: Basic Books.

SHEDLER, J., & BLOCK, J. (1980). Adolescent drug use and psychological health: A longitudinal inquiry. *American Psychologist* 45:612–630.

——— & WESTEN, D. (1998). Refining the measurement of Axis II: A Q-sort procedure for assessing personality pathology. *Assessment* 5:335–355.

TASK FORCE ON PROMOTION AND DISSEMINATION OF PSYCHOLOGICAL PROCEDURES (1995). Training and dissemination of empirically validated treatments. *The Clinical Psychologist* 48:3–23.

VANDENBOS, G.R. (1996). Outcome Assessment of Psychotherapy: Special Issue. *American Psychologist* 51.

WESTEN, D., & ARKOWITZ-WESTEN, L. (1998). Limitations of Axis II in diagnosing personality pathology in clinical practice. *American Journal of Psychiatry* 155:1767–1771.

——— & SHEDLER, J. (1999a). Revising and assessing Axis II: I. Developing a clinically and empirically valid assessment method. *American Journal of Psychiatry* 156:258–272.

——— & ——— (1999b). Revising and assessing Axis II: II. Toward an empirically based and clinically useful classification of personality disorders. *American Journal of Psychiatry* 156:273–285.

ZIMMERMAN, M. (1994). Diagnosing personality disorders: A review of issues and research methods. *Archives of General Psychiatry* 51:225–245.

Graduate School of Professional Psychology
University of Denver
2450 South Vine Street
Denver, CO 80208
E-mail: Jonathan@shedler.com