

Diagnosis of Personality Disorder: A Radical Behavioral Alternative

Kelly Koerner, Robert J. Kohlenberg, and Chauncey R. Parker
University of Washington

From a radical behavioral perspective, a single diagnostic system is unlikely to serve diverse purposes. Instead, ideal diagnostic systems should be developed to serve specific purposes. For example, the crucial proof required of a classification system designed to improve the outcome of psychosocial intervention would be that it enhance the clinician's influence on processes associated with client change. This means, in turn, that the change processes must be known or theoretically specified. As an illustration of this general approach to diagnostic classification, a specific behavioral theory is used to articulate processes of change in psychotherapy. The Axis II of the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; American Psychiatric Association, 1994) is then evaluated with respect to its ability to enhance clinicians' influence of these processes, found problematic, and an alternative classification criterion is proposed.

In many ways, the title of this article is self-contradictory. Radical behaviorists generally do not diagnose (in the tradition of the *Diagnostic and Statistical Manual of Mental Disorders*; e.g., 4th ed.; *DSM-IV*; American Psychiatric Association 1994), nor do they typically use the term *disorder*. Furthermore, they reject the notion of a personality "structure." This perhaps accounts for the relative absence of writings on personality disorder from a radical behavioral perspective. However, contemporary behaviorists use behaviorism as a means of understanding and promoting integration of methods and concepts used in other theoretical systems (Jacobson, 1991, 1992; Kohlenberg, Tsai, & Dougher, 1993), and a leading behavioral treatment of borderline personality disorder has distinct radical behavioral overtones (Linehan, 1993). Radical behaviorism's outsider position has the potential to yield a different view of the topic of diagnosis and personality disorders and perhaps provide some insights and solutions to problems. In this article, we use a radical behavioral approach to understand diagnosis and to specify the essential criteria of an ideal diagnostic system. We then evaluate concepts used in Axis II classification in light of what we view as the essential purpose of classification and offer suggestions for improving the system.

What Is Diagnosis?

Diagnosing or classifying (whether it be of personality disorder or any other type of disorder) is behavior, that is, something that mental health professionals and researchers do. According to behavioral theory, the key to understanding diagnosing (or any other behavior for that matter) is to specify its purpose (in behavioral terms, its "function"). The prevailing view seems to be that a single diagnostic system could be developed that would

be useful for multiple purposes such as making treatment, administrative, and legal decisions; facilitating communication with other clinicians and researchers; and guiding scientific inquiry. We see no a priori reason to assume that any classification system could be designed to adequately serve these differing functions. Those interested in treatment utility, forensics, psychopathology research, or cross-cultural issues have sufficiently different purposes that the same classification system will not serve for each. For example, if one is interested in heritable aspects of psychopathology, an optimal system of classification may specify latent variables to do with temperament; if one is interested in response to pharmacotherapy, then it might be optimal to classify according to biological correlates or similar pathophysiology. If one is interested in psychosocial intervention, then adequate classification would be based on the processes associated with psychological and social change. We see no reason to assume that the method of classification most useful for legal decisions would be the best for pharmacological or psychosocial intervention.

The belief that a single classification system could be developed to serve diverse purposes relies on the premise that there is a common underlying disorder or entity that explains a wide range of problematic behavior. It is this hypothetical entity to which the forensic psychologist, clinician, and psychopathology researcher can refer. Our behavioral orientation leads us, at the most fundamental level, to reject that premise and instead rely on explanations in terms of the problem behavior's context. This contextual view is consistent with Cloninger: "The structure of the observed behavioral variation may not be the same as the structure of the underlying biological variation because social learning and environmental factors also influence behavior" (1987, p. 574). Said differently, reductionism will not unambiguously solve classification problems. All needs are unlikely to be met with the same system.

Selecting a Purpose

We argue that classification systems should be developed according to the purpose for which they are meant—if the end

Kelly Koerner, Robert J. Kohlenberg, and Chauncey R. Parker, Department of Psychology, University of Washington.

Correspondence concerning this article should be addressed to Robert J. Kohlenberg, Department of Psychology, Box 351635, University of Washington, Seattle, Washington 98195-1635.

products overlap, it would be both a pleasant surprise and of considerable theoretical interest. In this article, we focus our analysis of the diagnosis of personality disorder for the purpose of improving the outcome of psychotherapy. From a behavioral perspective, this means that a diagnostic system should be devised and evaluated with respect to the goal of increasing the clinician's influence on processes associated with client change. This, in turn, requires that the essential elements of the change process be known or theoretically specified. In other words, we define an optimal diagnostic system as one that links problems, outcomes, and proposed processes of change. This position is consistent with the principle of problem-treatment-outcome (P-T-O) congruence suggested for psychotherapy research.

The principle of P-T-O congruence proposes that the intelligibility of psychotherapy research is a function of the similarity, isomorphism, or congruence among how we conceptualize and measure the clinical problem (P), the processes of therapeutic change (T), and the clinical outcome (O). (Strupp, Schacht, & Henry, 1988, p. 7)

Partly because of an absence of P-T-O congruence, the field lacks definitive empirical evidence on the processes of change in psychotherapy that could form the basis of a classification system for enhancing treatment utility. In the absence of data, it is necessary to select a theory of change that can inform diagnostic classification. In this article, we use a behavioral theory of psychotherapeutic change, known as functional analytic psychotherapy (FAP; Kohlenberg & Tsai, 1991), to illustrate the more general notion that the diagnostic process should help clinicians affect change processes. However, the general functional approach that we propose (that the diagnostic process should help the clinician affect the theorized mechanisms of change) would work with other behavioral and nonbehavioral theories as well.

How Therapy Works

The behavioral theory on which our approach to treatment is based is deceptively simple—individuals act the way they do because of the contingencies of reinforcement they have experienced in past relationships and, more broadly, in their culture. On the basis of this theory, it follows that clinical improvements or psychotherapeutic change also involves contingencies of reinforcement that occur in the relationship between client and therapist. A well-known aspect of reinforcement is that the closer in time and place the behavior is to its consequences, the greater the effect of those consequences. Treatment effects will be stronger, therefore, if clients' problem behaviors and improvements occur during the session, where they are closest in time and place to the available reinforcement. For example, if a client has difficulty trusting others, the therapy will be much more powerful if he experiences distrust in the therapeutic relationship where the therapist can respond immediately, as opposed to talking about experiences of distrust that occurred between sessions. From this view, therapeutic change results from the interpersonal contingencies that occur during the therapy session within the client-therapist relationship.

This theory does not say that other types of interventions (e.g., giving advice, using homework, cognitive therapy, social

skills training, or even free association as discussed later) are ineffective. The position is that in vivo interventions are more powerful and therapy outcome is improved if the therapist makes use of opportunities for intervening in this way. For example, Kohlenberg and Tsai (1994b) enhanced the effectiveness of conventional cognitive therapy for depression by adding an in vivo component suggested by FAP (Kohlenberg & Tsai, 1991).

According to this theory of change, improved therapeutic outcome depends on three interrelated factors. First, the client's daily life problems must also occur during the session within the context of the therapist-client relationship. The therapist-client relationship often evokes many of the same problem behaviors that trouble the client in other relationships. In addition, the parameters and structure of therapy (e.g., asking for help, limited session time or number of sessions, fees) may evoke other relevant problem behaviors that occur outside of session. We refer to these behaviors as clinically relevant behaviors (CRB), and they resemble the psychoanalytic conception of transference (for a more detailed comparison between psychoanalysis and FAP, see Kohlenberg & Tsai, 1994a). We refer to two of the three types of CRB used in FAP. The first type of clinically relevant behavior is the daily life problems the client experiences that also occur within the session (CRB1). A second type of clinically relevant behavior is naturally occurring within-session improvement (CRB2).

This leads to the second factor on which therapeutic improvement depends. The therapist must be aware (notice, observe) of the occurrence of clinically relevant problems and improvements. As detailed by Kohlenberg and Tsai (1991), the noticing of CRB by the therapist has many therapeutic benefits, particularly because the process of noticing is a precursor of natural reinforcement of clinical improvement, CRB2. In other words, if one notices even slight improvement, one is more apt to respond to it appropriately. The third factor leading to therapeutic improvement is the therapist encouraging and responding to CRB2 in a way that strengthens positive change. To fully use opportunities to promote change in session, the therapist must differentiate clinical problems from clinical improvements and reinforce incremental improvements.

Several examples highlight the importance of the therapist's ability to recognize and respond appropriately to improvement that might go unnoticed or be mistaken for problem behavior. For example, for the client who rigidly adheres to agreements despite reasonable needs to change them, a request to adjust a fee when his financial status changes might indicate improvement (a CRB2), and a therapist response that takes that request as expected and reasonable may influence the patient to be more flexible. However, the same request from a patient who, as a rule, unreasonably expects others to meet his needs might be an example of exactly the type of behavior that creates problems outside the session (a CRB1) and, consequently, would require a different response to foster positive change. The therapist's ability to identify problems as they occur in session and to notice small but positive improvements allows more sensitive responding likely to strengthen client improvements.

Here the ideas of "shaping" and "context" are important. Shaping refers to successive approximations to some desired behavior. The behavior need not be simplistic. For example, a client might have the goal of becoming "more able to hold my own

and not lose myself" in intimate relationships. As a client is helped to move from "losing herself" to "holding her own," the therapist strengthens incremental steps toward desired behavior, taking into account the behaviors present and absent from the client's repertoire (e.g., client is assertive at work but has difficulty recognizing her own preferences in intimate relationships). If this client indicates, in a subtle manner, that the therapist has misunderstood some issue from the week (e.g., by very quietly clearing her throat while the therapist finishes speaking), a naturally reinforcing response may be to explore, in a genuine manner, how the summary missed the mark and to subsequently better understand and respect the patient's point of view. Such a therapist response is likely to increase the probability that the patient will assert her views. To notice this slight improvement (expressing even a minimal negative reaction) is an example of naturally occurring change that is naturally reinforced when the clinician is sensitive to its occurrence.

Context refers to all of the circumstances that give a particular behavior meaning. From a behavioral perspective, this is phrased in terms of learning history and the contingencies that created and maintain behavior patterns. The context determines what constitutes a clinically relevant problem and distinguishes improvement from problematic behavior. The adaptive or maladaptive quality of a behavior depends on knowing the context in which the behavior occurred. For example, behavior that is topographically the same (arriving late to a session) might be a problem for someone who avoids (CRB1), but an improvement (CRB2) for someone who is compulsively on time. In this contextual approach to understanding people, a client's reinforcement history, environment, and circumstances are needed to give meaning to a particular behavior. Defining behaviors as pathological is not meaningful independent of context.

An optimal diagnostic system would help the clinician affect the processes associated with client improvement. The theory of psychotherapeutic change that we are advocating says that enhancing the therapist's ability to notice occurrences of in-session problems and improvements will almost inevitably result in better outcomes. From this perspective, an optimal diagnostic system would improve the therapist's ability to recognize clinically relevant behavior and would highlight the aspects of the context that help the therapist discern whether behavior is problematic or an improvement. Having sketched briefly the basic aspects of behavioral theory of psychotherapeutic change, we now evaluate the adequacy of the current Axis II diagnostic system for the purpose of improving the clinician's influence on processes of change.

An Evaluation of Concepts Used in Axis II

On the basis of our theory of change, the process of categorizing clients according to the *DSM-IV* is useful if it helps clinicians notice and respond appropriately to CRB. We now use the yardstick of treatment utility to assess the adequacy of the *DSM-IV* (American Psychiatric Association, 1994) Axis II constructs (traits and groupings of traits into personality disorder categories).

The *DSM-IV* defines traits as "enduring patterns of perceiving, relating to, and thinking about the environment and oneself

that are exhibited in a wide range of social and personal contexts" (American Psychiatric Association, 1994, p. 630). As currently specified, the *DSM-IV* directs the diagnostician to pay attention to personality traits under certain conditions. According to the *DSM-IV*:

Only when personality traits are inflexible and maladaptive and cause significant functional impairment or subjective distress do they constitute Personality Disorders. The essential feature of a Personality Disorder is an enduring pattern of inner experience and behavior that deviates markedly from the expectations of the individual's culture and is manifested in at least two of the following areas: cognition, affectivity, interpersonal functioning, or impulse control (Criterion A). This enduring pattern is inflexible and pervasive across a broad range of personal and social situations (Criterion B) and leads to clinically significant distress or impairment in social, occupational, or other important areas of functioning (Criterion C). The pattern is stable and of long duration, and its onset can be traced back at least to adolescence or early adulthood (Criterion D). The pattern is not better accounted for as a manifestation or consequence of another mental disorder (Criterion E) and is not due to the direct physiological effects of a substance . . . or a general medical condition . . . (Criterion F). (American Psychiatric Association, 1994, p. 630)

Traits on Axis II are grouped into specific personality disorder (PD) categories (i.e., schizoid, histrionic, avoidant, etc.) as well as into descriptive clusters (A: odd/eccentric, B: dramatic/emotional/erratic, and C: anxious/fearful). Diagnostic categories can be viewed as a set of covarying responses. For a given category to be useful, an individual should have a high probability of showing the diagnostic criteria of that category (Nelson & Barlow, 1981) and have a low probability of showing diagnostic criteria from other categories. In the best case, diagnostic groupings clearly describe which behaviors covary and, therefore, could be very useful to clinicians in noticing CRB. For example, if the client does *x* during the session, it then alerts the therapist to watch for *y*: If a client rapidly shifts among emotions and expresses shallow emotion, the therapist might watch for ways the client is easily influenced by therapist comments; if the client has a sense of entitlement, the therapist might watch for ways the client expects admiration and envies the therapist or believes the therapist envies him. Description of behavior that hangs together is a desirable goal of a diagnostic system, because it enhances therapist sensitivity to in-session behavior.

Unfortunately, current distinctions between Axis I and Axis II and among Axis II disorders do not serve that purpose well. First, there is significant overlap of Axis I and II (Livesley, Schroeder, Jackson, & Jang, 1994). For example, a variety of anxiety disorders overlap with avoidant and dependent PD and traits (Taylor & Livesley, 1995). PD categories also overlap with each other. Individuals often meet criteria for more than one PD (e.g., Clarkin, Widiger, Frances, Hurt, & Gilmore, 1983; Kass, Skodol, Charles, Spitzer, & Williams, 1985) and these overlaps are not random (e.g., Widiger, 1992; Widiger, Frances, Warner, & Bluhm, 1986). Furthermore, evidence of the validity of cluster distinctions is mixed (e.g., Bell & Jackson, 1992; Dowson & Berrios, 1991; Hyler et al., 1990; Kass et al., 1985; Livesley, Jackson, & Schroeder, 1989, 1992; Morey, 1988; Pilonis & Frank, 1988). The bottom line, then, is that current classifications on Axis II offer limited help in predicting which

behaviors will co-occur. The current grouping of traits into specific PDs, into clusters, or into Axis I versus Axis II does not give clinicians adequate guidance to notice and respond appropriately to in-session behavior. Thus, according to our behaviorally derived yardstick, the *DSM-IV* doesn't measure up.

Why the *DSM-IV* Does Not Work

The problem, in our view, is fundamental. When the purpose of classification is to enhance the clinician's ability to identify functional relationships, a trait-based diagnostic system is problematic. The behavioral theory of psychotherapeutic change puts a premium on identifying controlling variables (i.e., the conditions that give rise to and maintain problem behaviors and improvements). Traits, although descriptive, tell nothing about controlling variables. Take for example, the trait of aggression or hostility. If classified only by topography, the function of behavior is missed. Consider two individuals who experience and express intense, problematic anger, including physical aggression toward others. One's hostility/aggression is primarily communicative and maintained by the effect of extreme emotional displays on others, whereas the other's hostility/aggression is primarily emotional avoidance, a classically conditioned secondary emotional response evoked in situations that elicit sadness. In the first situation, contingency management is called for; in the second, exposure and response prevention. It is not clear that trait concepts alone would lead to assessment of controlling variables that indicate treatment interventions.

Said differently, trait concepts miss variability that is interesting and clinically useful. For example, Benjamin and Wonderlich (1994) describe how trait measurement of aggression missed the variability of interest that was detected by a measure including more interpersonal context. Similarly, Paulhus and Martin (1987) discuss how treating hostility and nurturance as opposite traits, rather than as abilities that change according to the setting, ignores that "individuals capable of hostility are also capable of nurturance" (p. 354).

Even worse, trait concepts may at times interfere with assessing the function of behavior. When behavior is described topographically and independent of context or when behavior is taken as a sign of a latent entity, traits can become easily reified into causal entities that give the feeling of having explained behavior but simultaneously derail further explanation needed for clinical intervention. A view of traits as entities that take on causal agency can divert the clinician from assessing the circumstances under which hostile behavior does and does not occur.

We prefer that the cross-situational stability to which traits refer be discussed in terms of processes that have a high probability across situations, rather than as causal entities or structures. We use the term *cross-situational response classes*, or simply "response classes." Response classes are hypothetical groupings of behavior that serve the same function for an individual. Behaviors are grouped into a response class when they are maintained by the same discriminative and reinforcing stimuli. This distinction is more than semantic in that these competing views have different effects on a clinician's behavior. We have argued that a diagnostic system should promote effective identification of CRB and sensitivity to subtle change. The behaviors that are considered to indicate a trait may form

a useful response class, but a trait analysis need not extend the analysis to environmental determinants in the way that a functional analysis of response classes would. For example, explaining a lengthy angry silence as expectable because clients who were diagnosed as having borderline PD have more trait anger than those who were not is less useful than an analysis of the conditions in session that provoke and maintain the behavior.

For the radical behaviorist, traits as entities are antithetical to a functional analysis, and we have concluded that there appears to be little treatment utility derived from holding a trait view as detailed in the *DSM-IV*. Nevertheless, we think it is useful to consider the potential benefits that a structural or entity view of traits may hold. From our behavioral viewpoint, it is possible to have practical benefits derived from a therapist believing in the trait view, although we reject the notion on a conceptual level. Again, we use our behaviorally based yardstick to assess the potential benefits of a therapist holding a trait view.

Using trait concepts from the *DSM-IV* Axis II may, in some cases, alert the therapist to problems and improvements as they occur in session. First, knowing the behaviors likely to be included by, for example, trait anxiety, can help the clinician notice seemingly disparate behaviors that serve a similar purpose. To the extent that a trait is defined by behaviors that do form a response class, the trait name can be a helpful shorthand label. Second, research on personality traits emphasizes the continuity of normal and abnormal behavior. Research specific to PD suggests that the distinction between normal and abnormal personality is arbitrary (Frances, 1980) and that traits are continuously distributed with similar intercorrelations among traits in individuals with and those without PD (Livesley et al., 1992; Tyrer & Alexander, 1979). Therapists influenced by this data might be less likely to stigmatize behavior or misconstrue behavior as pathological. In our terms they would be more sensitive to the contextual nature of disorder and likely to reinforce improvement (CRB2). Remembering that the distinction between normal and abnormal personality is arbitrary might help the therapist note improvement because a continuum of health to disorder is implied (cf. Follette, Bach, & Follette, 1993). Third, to the extent that Axis II is synonymous with more difficult-to-treat problems, a diagnostic process that identifies those problems could help clinicians be more tolerant of slow change and maintain a naturally reinforcing stance rather than become impatient and fail to reinforce CRB2. Fourth, research suggests that a relatively small number of behavioral dimensions appear to capture Axis II (Grove & Tellegen, 1991). For example Livesley, Jang, Jackson, & Vernon (1993) found four factors: lability, antagonism, social inhibition, and compulsivity. This, therefore, reduces the complexity of having to consider all possibly CRB to a smaller set that are likely to be relevant to Axis II. Nevertheless, on the basis of the theory of psychotherapeutic change that we are proposing, a diagnostic system should foster therapist sensitivity to contingencies. Traits can, but need not, overlap with response classes and, at times, discourage analysis of contingencies.

Toward the Goal of Enhancing Attention to CRB

We have argued that a diagnostic system intended to enhance psychotherapy outcome should be explicitly tied to a theory of

change. The results of making a diagnosis should enhance the clinician's impact on processes of change. Consistent with a behavioral theory of change, a classification system would point to contingencies and response classes rather than traits.

A second step in keeping with this theory of change is to add the diagnostic requirement that the client's daily life problems occur in session. We would retain the *DSM-IV* criteria that behavior be enduring, cross-situational, cause distress, and so forth, but further require that the behavioral patterns must also occur in session. This would be a marked departure from current diagnosis, but we believe this is most likely to promote changes in the therapist that will, in turn, facilitate psychotherapeutic change. Requiring that the behavior happen in session as well as outside session would underscore the importance of observing in-session problematic behavior and using in vivo treatment.

In this system, clients would be classified according to the specific stimulus functions and response classes they bring to therapy. The system would foster assessment of functional relationships of interpersonal behavior patterns that occur in the client's daily life and during the session. The system would specify stimulus control, response repertoires, and restricted or inappropriate reinforcement stimuli associated with problematic interpersonal behavior.

Although some current Axis II items imply a functional relationship between interpersonal problems and controlling variables, they do not make these relationships explicit enough for our purpose. For example, "suspects, without sufficient basis, that others are exploiting, harming, or deceiving him or her" indicates that the individual may have problems with appropriate discrimination in social situations. "Shows restraint within intimate relationships because of the fear of being shamed or ridiculed" indicates the possibility of conditioned emotional responding and avoidance/escape. "Has difficulty making everyday decisions without an excessive amount of advice and reassurance from others" suggests deficits in the behavioral repertoire and possible contingencies that maintain problem behavior. Items like these help orient the therapist to relevant problem behaviors that may occur in session. A diagnostic system consistent with a behavioral theory of psychotherapeutic change, however, would make these functional relationships explicit.

On the basis of a functional analysis, Kohlenberg and Tsai (1991) viewed the therapist's private reactions and feelings as central to diagnosis, and treatment focused on within-session occurrences of the client's problematic behavior. In an ideal system, the therapist's reactions to the client's problematic behavior would be more systematically included as a diagnostic tool and as diagnostic criteria. For example, Pilkonis (personal communication, July 1995) describes typical (although not universal) reactions in therapists and others to particular behavioral patterns associated with Axis II problems. He has created a checklist for trainees of reactions elicited in the therapist by clients who meet criteria for various PD categories. For example, if the therapist feels remote, distanced, out of touch, disconnected, the patient is likely to have behavioral patterns listed under schizoid PD; if one feels demeaned and belittled, the patient may have aspects of narcissistic PD; if one feels inhibited, not wanting to wound or expose deficiencies, the patient may

have characteristics listed under avoidant PD. Clinicians frequently note the occurrence of in-session problems by observing their own reactions, and more generally, clinicians use their own emotional reactions to assess the stimulus functions of the client. Clinicians would be better able to identify relevant in-session behavior if a diagnostic system was informed by systematic study of the functional relationship of clinicians' reactions to client behaviors.

A Behavioral Approach to Problems of the Self

Our functional analytic approach emphasizes the role of interpersonal contingencies in the cause, treatment, and hence in the construction of a classification system for the behavioral phenomena known as PDs. At the same time, we question the utility of notions such as traits and entities. Because the self is typically treated as a nonbehavioral entity or trait, it might appear that our approach excludes or does not deal with the self or problems of the self that are a central feature in the definition of PDs. Appearances notwithstanding, our functional analysis does include an account of self (and problems of self) that seems to retain much of the essence of what is usually meant by self in discussions of PDs and yet avoids reference to entities or other mental constructs. The advantage to our approach is that it points to the kinds of variables that can be used in the design of a diagnostic system and treatment procedures.

The behavioral account of self presented here is a summary of a more detailed discussion by Kohlenberg and Tsai (1991, Chap. 6; 1995). Consistent with the *DSM-IV* description of PD as being an enduring pattern of "inner experience" (American Psychiatric Association, 1994, p. 633), the focus of our analysis is on the experience of self. In essence, we explain the experience of self by identifying and describing the something (the stimulus or object) that is experienced. Our analysis parallels that of explaining the experience of other more easily described objects (stimuli) such as heat, apples, and ice cream cones. To be sure, the stimulus that is the basis of the experience of self is not a typical object, in that it is supposed to be private, inside the person, and not visible to others. We say "supposed to be private" because, as we explain later, the conditions under which we acquire the self-experience can often lead to a mixture of both public and private origins of the relevant stimulus. It is this mixture of public and private elements that accounts for problems of the self often referred to in the *DSM-IV*.

There are seven important features of our view of self. First, we assume that much of the self experience is learned early in life, beginning at about the time a child learns to talk.

Second, we propose that the experience of self parallels the development of the child's use and personal meaning for self referents such as "I," "me," "Jamie," and "Chauncey" (in referring to self). The notion that there is a link between the experience of self and language development has been suggested and researched by nonbehaviorists throughout the years (Cooley, 1908; Fraiberg, 1977) and is consistent with behavioral conceptions of language.

Third, the personal meaning and use of self-referents are naturalistically taught by parents (or caretakers) by prompting, modeling, and reinforcing responses that they judge as appropriate (or punishing or ignoring inappropriate responses).

Fourth, unlike common objects such as apples, self-referents cannot be taught directly but are learned indirectly from their use in sentences such as "I see Mommy" and "I am hungry" through a concept learning type process referred to by Skinner as the "by-product of the acquisition of larger responses containing identical elements" (1957, p. 120). We refer to this learning process as emergence.

Fifth, there is a developmental progression in the emergence of "I" (although we refer to "I," our analysis applies to all self-referents) as an independent unit and correspondingly in the self-experience. In Stage 1, the parent teaches the toddler concrete phrases such as "I want juice, I want ice cream, I want doggy," as well as numerous other similarly structured phrases such as "I am hot, I am hungry, I am here"; "I see car, I see mommy"; and "I feel sad, I feel happy, I feel icky." Stage 2 occurs after the child has had enough exemplars of concrete phrases, and the more abstract utterances such as "I want," "I am," "I see," and "I feel" emerge and have independent meaning. It is at this point that the child can say "I want candy," although he or she has never combined "I want" and "candy" before. In Stage 3, the even more abstract term "I" emerges and has independent meaning. The progression from larger to smaller verbal units described earlier has been observed in children as they learn to talk (Cooley, 1908; Dore, 1985; Fraiberg, 1977; Peters, 1983).

Sixth, the nature of the stimulus that corresponds to the independent "I" is derived from the stimuli used for phrases in the previously learned stage. In Stage 1, the stimuli used by the parent are almost completely public. For example, the parent might prompt the child to say "baby wants juice" if the baby is publicly observed to be looking at or reaching for the juice. If the parent is correct, there are also private stimuli present for the child (e.g., an internal "wanting of juice"). After a number of such exemplars involving the parent being correct, the private stimulus of "wanting" is present (regardless of what is wanted) and comes to be the stimulus that evokes the Stage 2 "I want." Similarly, in Stage 3, an independent "I" emerges after there have been many exemplars from Stage 2 involving private stimuli being present for "I feel," "I see," and so forth. We describe the private stimulus on which "I" is based as perspective (after Hayes, 1984). Perspective becomes the relevant stimulus for "I" because it is the only element that is present across all situations containing "I," as in "I am," "I have," "I want," and "I see." In normal development, this private, relevant stimulus (perspective) is unique in that it is always where the child is and does not change as the child's body changes (e.g., as the child grows, has chicken pox, etc.). It is also the place (location) where the wanting occurred that controls the "I want," as well as the "seeing," "feeling," and so forth for the various "I" (verb forms)." These properties of the relevant stimuli that control "I" account for the everyday nonpathological experience of "I" as "an abiding, resting awareness, featureless and unchanging, a central something that is witness to all events, exterior and interior" (Deikman, 1973, p. 325) and might even be responsible for spiritual experience (Hayes, 1984). The evocation of "I" by the stimulus perspective gives rise to an individual who "knows" what they want, see, feel, and so forth, and these experiences are not unduly influenced by external stimuli such as

demands or expectations from other people. In a word, the individual has a consistent and unambiguous self-experience.

Seventh, the process can go awry if the parent is incorrect and does not interact with the child in ways that foster the eventual control of "I" by the private stimulus. In the example "baby want juice" discussed earlier, if the child is really reaching toward another object or is trying to push the juice away, then there is no private "wanting of juice" stimulus and one of the building blocks that form the foundation for private stimulus evocation of "I want." The process can go awry in Stage 2 if the parent doesn't reinforce (or even ridicules) a child who says "I see fishie" when this phrase has never been used before and is a composition by the child of "I see" and "fish" and the fish can only be seen by the child and not the parent. There is a continuum of private control of "I" that corresponds to the number of exemplars that were not brought under private control. Parents who are distracted, anxious, schizophrenic, and so forth are most prone to interfere with the development of private stimulus evocation of "I." In contrast to the stable sense of self described earlier, a continuum of self problems are the result of more or less private control of "I." At the mild end of the continuum, a somewhat unstable self occurs because it varies with who is present and in particular is variable in close relationships. This occurs because public control over the "I want, I see, I am (e.g., hungry)," largely involves the parent, an important, close relationship. At the other end of the continuum, the sense of self is largely influenced by external stimuli such as the opinions, beliefs, desires, and moods of others. Under these conditions, life is unpredictable, chaotic, and subject to the whims of others, which could result in the chaotic interpersonal relating and emotionality of borderline PD.

Implications for a Diagnostic System

Our analysis of self leads to suggestions for methods for the in-session assessment of problems of the self. In general terms, our theory of self would direct the therapist to be vigilant for CRB such as a client acting wary, overly attentive, and concerned with the therapist's opinions. A more systematic assessment would involve using a probe that has more or less structure to determine the degree of external control over self. For example, Kohlenberg and Tsai (1991, Chap. 6) suggest using variants of free association for this purpose and report on its clinical effects. The free association task can be presented with more or less structure. When used in this manner, the primary purpose of free association is not to uncover hidden meanings or to make use of the content in the associations. Instead, it is the behavior of free associating that is of interest. In its most unstructured form, the free association instructions are as follows: "Tell me everything that enters your mind—all thoughts, feelings, and images. It's important not to censor anything. Report whatever comes up, even if you think it's unimportant, nonsensical, trivial, embarrassing, or whatever." The client is asked to continue to do this without feedback from the therapist and may even be asked to sit so that the therapist is out of view.

Our view of this task is that it will evoke self-experiential problems for individuals whose "I" experience is dependent on external stimuli. The task requires talking to another person (the therapist) with a minimum of external cues from the lis-

tener. In essence, the instructions are “give me ‘I’ statements and observe your inner self based on private stimulus control—that which is purely inside you and has nothing to do with me, my facial expressions, feelings, etc.” Under these conditions, it is possible for the client to say “I feel x (where x is any feeling),” or “I see this (particular image)” under conditions that favor control by private stimuli. On the basis of Kohlenberg and Tsai’s case material, clients with extensive self problems became very anxious and were unable to perform this task because of the lack of public stimulation. As illustrated in the following example, clients seemed to experience a “loss of self” in the absence of therapist cues. Variations in the classical unstructured format of free association were used to assess the degree of public control versus private control of the self experience. For example, the following is a transcript of comments made by Fred (a client classified as avoidant personality disorder) when first asked to do open-ended, free association with his eyes closed.

O.K. [A long, long pause]. Terrible [anxious laughter]. . . I, I just can’t [A long pause]. I mean, it’s, I mean, I can’t, I can’t focus, it’s really embarrassing, you know, you ought to be able to do it. . . I mean, it’s just, it’s just like nothing, you know, I mean.

Fred then went on to describe how difficult the task was:

I think what’s happening there is, uh, I have to be able to back off a little bit, I mean, and I even try to do that and I’m having some trouble. . . just [pause], you know, being an observer in this situation. . . I think it would just simply get worse. The longer it went on.

On the basis of Fred’s reaction, the free association task was altered from an open-ended format to one of limited duration. He was given a 2-min, a 1-min, and a 15-s version of free association. He was able to report “I” statements pertaining to his experience during the 15-s version but not at longer durations. The reported difficulty and anxiety appeared to be a function of the duration of the task.

This type of assessment leads to the therapist attending to in-session instances of self problems. Consistent with our treatment utility view of classification, it then leads the therapist to in-session treatment aimed at the development of a sense of self that is less dependent on others. In Fred’s case, for example, the treatment included gradually increasing the duration of free association. In addition to using the graded series of free association tasks, the therapist would provide consistent reinforcement for private control by taking seriously and being responsive to statements about the self (“I want,” “I feel,” “I am”) that are under control of private states rather than of the therapist, encouraging the client to talk about private states in the absence of specific external cues, and so forth.

Conclusion

We have discussed a diagnostic system based on interpersonal problems that occur in session in the context of the therapeutic relationship. Because our approach includes problems of the self as both interpersonal and occurring within session, our diagnostic system would probably overlap significantly with PD according to *DSM-IV*. We would also expect differences as well. Clearly, the specifics of the alternative we are proposing are

undeveloped. Furthermore, we realize that the alternatives that we are proposing raise many methodological problems. First, not all problems are interpersonal, have to do with the self, or occur in session. Furthermore, if classification for psychotherapy moved in the direction that we propose, diagnosis would depend on the therapist’s ability to detect and evoke CRB in session. However, if we are to remain consistent with our approach, then these problems must be tackled to develop adequate diagnostic classification.

Second, issues of measurement and reliability are raised. On one hand, including a within-session requirement for diagnosis means that the behavior can be directly observed (even videotaped), thus allowing for consensus and more precise definition of what is meant (i.e., one could show the tape to illustrate what is meant by “anger,” “withdrawal,” “paranoia,” or “loss of self”). Direct observation might help to remedy reliability problems (Livesley & Jackson, 1992). On the other hand, the increased reliance on the therapist sophistication may not necessarily increase reliability. Although we expect that there would be some meaningful commonality across individuals’ response classes, a functional analysis is so different than current diagnostic practice that it is hard to detail exactly how it may diverge from current Axis II.

Furthermore, an adequate functional assessment may require that methods of assessment be tailored to the problem—self-report, collateral informants, clinical observation, will vary in their usefulness to identify the problematic behavior and the conditions that maintain it. In other words, there may be a need for problem-specific assessment guided by the premise that the behavior is a sample of the relevant clinical behavior, rather than a sign of a latent construct.

These methodological challenges are surmountable. Guided by concern for good measurement, a method of assessing in-session behavior could have adequate and sufficient sampling of behavior that is reasonably free from measurement bias and is representative of the patient’s real-life problems. These response classes could be subjected to conceptual and empirical investigation to develop a diagnostic system that would improve the clinician’s attention to in-session problems and improvements. Ignoring for a moment the particulars of the behavioral theory of change and its implications for diagnostic classification, what we have tried to illustrate more generally is the utility of developing classification that is meant to serve a single clearly stated purpose. We believe that it is unlikely that refinements to the *DSM* will improve treatment utility unless a much more fundamental revision linked to change processes is undertaken, from whatever theoretical perspective. We hope that the field will move toward alternative classification systems that will better serve the purpose of improving the outcome of psychotherapy.

References

- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- Bell, R. C., & Jackson, H. J. (1992). The structure of personality disorders in *DSM-III*. *Acta Psychiatrica Scandinavica*, 85, 279–287.
- Benjamin, L. S., & Wonderlich, S. A. (1994). Social perceptions and borderline personality disorder: The relation to mood disorders. *Journal of Abnormal Psychology*, 103, 610–624.

- Clarkin, J. F., Widiger, T. A., Frances, A., Hurt, S. W., & Gilmore, M. (1983). Prototypic typology and the borderline personality disorder. *Journal of Abnormal Psychology, 92*, 263-275.
- Cloninger, C. R. (1987). A systematic method for clinical description and classification of personality variants: A proposal. *Archives of General Psychiatry, 44*, 573-588.
- Cooley, C. H. (1908). A study of the early use of self words by children. *Psychological Review, 15*, 339-357.
- Deikman, A. J. (1973). The meaning of everything. In R. E. Ornstein (Ed.), *The nature of human consciousness*. San Francisco: Freeman.
- Dore, J. (1985). Holophrases revisited: Their "logical" development from dialogue. In M. Barret (Ed.), *Children's single word speech*. New York: Wiley.
- Dowson, J. H., & Berrios, G. E. (1991). Factor structure of *DSM-III-R* personality disorders shown by self-report questionnaire: Implications for classifying and assessing personality disorders. *Acta Psychiatrica Scandinavica, 84*, 555-560.
- Follette, W. C., Bach, P. A., & Follette, V. M. (1993). A behavior-analytic view of psychological health. *Behavior Analyst, 16*, 303-316.
- Fraiberg, S. (1977). *Insights from the blind. Comparative studies of blind and sighted infants*. New York: Basic Books.
- Frances, A. J. (1980). The *DSM-III* personality disorders section: A commentary. *American Journal of Psychiatry, 137*, 516-527.
- Grove, W. M., & Tellegen, A. (1991). Problems in the classification of personality disorders. *Journal of Personality Disorders, 5*, 31-41.
- Hayes, S. C. (1984). Making sense of spirituality. *Behaviorism, 12*, 99-110.
- Hyler, S., Lyons, M., Rieder, R. O., Young, L., Williams, J. B. W., & Spitzer, R. L. (1990). The factor structure of self-report *DSM-III* Axis II symptoms and their relationship to clinicians' ratings. *American Journal of Psychiatry, 147*, 751-757.
- Jacobson, N. S. (1991). *Presidential address*. Presented at the annual meeting of the Association for the Advancement of Behavior Therapy, Boston.
- Jacobson, N. S. (1992). Behavioral couple therapy: A new beginning. *Behavior Therapy, 23*, 493-506.
- Kass, F., Skodol, A. E., Charles, E., Spitzer, R. L., & Williams, J. B. W. (1985). Scaled ratings of *DSM-III* personality disorders. *American Journal of Psychiatry, 142*, 627-630.
- Kohlenberg, R. J., & Tsai, M. (1991). *Functional analytic psychotherapy: Creating intense and curative therapeutic relationships*. New York: Plenum.
- Kohlenberg, R. J., & Tsai, M. T. (1994a). Functional analytic psychotherapy: A radical behavioral approach to treatment and integration. *Journal of Psychotherapy Integration, 4*, 175-201.
- Kohlenberg, R. J., & Tsai, M. (1994b). Improving cognitive therapy for depression with functional analytic psychotherapy: Theory and case study. *Behavior Analyst, 17*, 305-320.
- Kohlenberg, R. J., Tsai, M. (1995). I speak, therefore I am: A behavioral approach to understanding the self. *Behavior Therapist, 18*, 113-116.
- Kohlenberg, R. J., Tsai, M., & Dougher, M. J. (1993). The dimensions of clinical behavior analysis. *Behavior Analyst, 16*, 271-282.
- Linehan, M. (1993). *Cognitive-behavioral treatment of borderline personality disorder: The dialectics of effective treatment*. New York: Guilford Press.
- Livesley, W. J., & Jackson, D. N. (1992). Guidelines for developing, evaluating, and revising the classification of personality disorders. *Journal of Nervous and Mental Diseases, 180*, 609-618.
- Livesley, W. J., Jackson, D. N., & Schroeder, M. L. (1989). A study of the factorial structure of personality pathology. *Journal of Personality Disorders, 3*, 292-306.
- Livesley, W. J., Jackson, D. N., & Schroeder, M. L. (1992). Factorial structure of traits delineating personality disorders in clinical and general population samples. *Journal of Abnormal Psychology, 101*, 432-440.
- Livesley, W. J., Jang, K. L., Jackson, D. N., & Vernon, P. A. (1993). Genetic and environmental contributions to dimensions of personality disorder. *American Journal of Psychiatry, 150*, 1826-1831.
- Livesley, W. J., Schroeder, M. L., Jackson, D. N., & Jang, K. L. (1994). Categorical distinctions in the study of personality disorder: Implications for classification. *Journal of Abnormal Psychology, 103*, 6-17.
- Morey, L. C. (1988). The categorical representation of personality disorder: A cluster analysis of *DSM-III-R* personality features. *Journal of Abnormal Psychology, 97*, 314-321.
- Nelson, R. O., & Barlow, D. H. (1981). An overview of behavioral assessment with adult clients: Basic strategies and initial procedures. In D. H. Barlow (Ed.), *Behavioral assessment of adult disorders* (pp. 13-43). New York: Guilford Press.
- Paulhus, D. L., & Martin, C. L. (1987). The structure of personality capabilities. *Journal of Personality and Social Psychology, 52*, 354-365.
- Peters, A. N. (1983). *The units of language acquisition*. London: Cambridge University Press.
- Pilkonis, P. A., & Frank, E. (1988). Personality pathology in recurrent depression: Nature, prevalence, and relationship to treatment response. *American Journal of Psychiatry, 145*, 435-441.
- Skinner, B. F. (1957). *Verbal behavior*. New York: Appleton-Century-Crofts.
- Strupp, H. H., Schacht, T. E., & Henry, W. P. (1988). Problem-treatment-outcome congruency: A principle whose time has come. In H. Dahl, H. Kachele, & H. Thoma (Eds.), *Psychoanalytic process research strategies* (pp. 1-14). Berlin, Germany: Springer-Verlag.
- Taylor, S., & Livesley, W. J. (1995). The influence of personality on the clinical course of neurosis. *Current Opinion in Psychiatry, 8*, 93-97.
- Tyrer, P., & Alexander, J. (1979). Classification of personality disorder. *British Journal of Psychiatry, 135*, 163-167.
- Widiger, T. A. (1992). Categorical versus dimensional classification: Implications from and for research. *Journal of Personality Disorders, 6*, 287-300.
- Widiger, T. A., Frances, A., Warner, L., & Bluhm, C. (1986). Diagnostic criteria for the borderline and schizotypal personality disorders. *Journal of Abnormal Psychology, 95*, 43-51.

Received August 31, 1995

Revision received November 30, 1995

Accepted December 12, 1995 ■