



PII S0272-7358(96)00038-4

ASSESSMENT OF NARCISSISTIC PERSONALITY DISORDER: A MULTI-METHOD REVIEW

Mark J. Hilsenroth

University of Arkansas

Leonard Handler

University of Tennessee

Mark A. Blais

*Department of Psychiatry, Massachusetts General Hospital
and Harvard Medical School*

ABSTRACT. *This review examines the available empirical data for the diagnosis of narcissistic personality disorder (NPD) for three methods of assessment: semi-structured interviews, self-report inventories, and projective techniques. Issues of reliability, validity, and clinical utility are examined for each instrument (or scale). An overview of the relative advantages, disadvantages, and empirical support for each method of assessment in the diagnosis of NPD is presented in a discussion after the review of the salient literature. In general, it was found that semi-structured interviews are a fairly reliable and valid method of diagnosis for Axis II disorders but, for the most part, these studies have used woefully small samples of NPDs. In general, self-report instruments were best at screening for the presence or absence of personality disorder, identifying members of personality disorder clusters, and identifying negative instances of specific personality disorders or clusters. Self-report inventories and the structured interviews are often in disagreement concerning presence of specific personality pathology. In general, previous studies have found the tendency for self-report measures to diagnose personality disorders at much higher frequencies than do clinicians. Moreover, self-report measures frequently attributed two or more personality disorders to a particular individual. Additionally, research with projective methods over the last decade has shown this mode of assessment to be useful in the differential diagnosis of NPD from both*

Correspondence should be addressed to Dr. Mark J. Hilsenroth, 216 Memorial Hall, Department of Psychology, University of Arkansas, Fayetteville, AR 72701. E-Mail: hils@comp.uark.edu

related and unrelated personality disorders. It seems prudent that the clinician and researcher alike should employ multiple methods of measurement and utilize information in a systematic and theoretical fashion when evaluating a patient for NPD diagnosis. Copyright © 1996 Elsevier Science Ltd

NARCISSISTIC PERSONALITY DISORDER (NPD) was included as a diagnostic category in the *Diagnostic and Statistical Manual of Mental Disorders (3rd ed.) (DSM-III)* (American Psychiatric Association, 1980) in large part due to widespread interest in the theoretical and clinical concept of narcissism by psychotherapists (Kernberg, 1970, 1975; Kohut, 1971, 1977; Meissner, 1978; Pulver, 1970; Stolorow, 1975; Teicholz, 1978). However, justification for regarding NPD as an independent diagnostic entity having distinguishable features from other personality disorders has been the matter of some controversy (Loranger, Oldham, & Tulis, 1982; Pope, Jonas, Hudson, Cohen, & Gunderson, 1983; Masterson, 1981; Siever & Klar, 1986; Vaillant & Perry, 1985). In fact, almost no empirical work focusing exclusively on NPD had been conducted until the late 1980s. In their review of data concerning *DSM-III-R* (American Psychiatric Association, 1987) descriptors of NPD Gunderson, Ronningstam, and Smith (1991) stated that "it [NPD] remains a disorder about which there has been little empirical evidence and around which basic questions of description, clinical utility and validity still remain" (p. 167).

However, recent efforts have begun to systematize and describe the characteristic features of NPD. In this effort to illuminate the various questions surrounding the nosological aspects of NPD, investigators have attempted to develop specific criteria on a number of assessment measures that may aid in the differentiation of NPD from other personality disorders. In particular, several phenomenological studies by Gunderson and Ronningstam (Gunderson, Ronningstam, & Bodkin, 1990; Ronningstam, 1988; Ronningstam & Gunderson, 1988, 1990, 1991) have focused on identifying different characteristics of NPD patients and has led to the development of the *Diagnostic Interview for Narcissism* (DIN). Also, self-report measures have been shown to be useful in the diagnosis of personality disorders in general and scales designed to assess NPD have received extensive use (Chatham, Tibbals, & Harrington, 1993; Colligan, Morey, & Offord, 1994; Millon, 1987; Morey, Waugh, & Blashfield, 1985; Wink & Gough, 1990). In addition, several studies using projective methods of assessment have shown promise in the differential diagnosis of NPD (Berg, 1990; Berg, Packer, & Nunno, 1993; Farris, 1988; Gacono, Meloy, & Heaven, 1990; Gacono, Meloy, & Berg, 1992; Hilsenroth, Hibbard, Nash, & Handler, 1993; Hilsenroth, Fowler, Padawer, & Handler, in press).

Several authors have called for research concerning the differential diagnosis and treatment of individuals suffering from character pathology (Berg, 1983; Blatt & Lerner, 1983; Lerner, 1988; Kernberg, 1975; Westen, 1990). All have stressed the importance of careful diagnostic assessment of these individuals, especially utilizing psychological testing, for treatment planning, management of transference, and countertransference issues. The ability to distinguish narcissistic pathology specifically, and the assessment of personality disorders in general, would enable practicing clinicians to make more appropriate decisions in choosing treatment strategies for such persons. Identification of variables related to pathological expressions of narcissism is only a starting point. A further step will be identification of those features of NPD that are the most outstanding and important in differential diagnosis (Davis, Blashfield, & McElroy, 1993). The fact that narcissistic traits commonly occur in other,

nonnarcissistic personality disorders has led to a variety of opinions as to the clinical features and theoretical aspects of NPD (Pulver, 1970; Teicholz, 1978; Chessick, 1987; Kernberg, 1984; Millon 1981). Therefore, it is imperative that indices of narcissism be identified that can separate NPDs from related "dramatic" Cluster B personality disordered patients [Antisocial (ANPD), Borderline (BPD), and Histrionic (HPD)] as well as the differentiating unrelated (Cluster A and Cluster C) personality disorders.

Often, assessment occurs along one modality which limits application of results and restricts the extent to which data can be generalized to both clinical and research applications. In the psychological assessment of pathological narcissism, one needs to allow for problems of disorder differentiation, comorbidity, sampling from a range of different severity, symptom overlap, and complexity. The need for a multimodal assessment of pathology is not a new concept; it has been proposed previously by Leary (1957) as well as Rapaport, Gill, and Schafer (1968). Our evaluation of the literature, then, will attempt to develop a more coherent diagnostic understanding of NPD through use of the multimethod approach which utilizes a clinical interview, self-report, and projective data. The goal of this approach to the diagnosis of NPD is to connect both the surface (readily apparent in behaviors) and deeper (intrapsychic) manifestations of this disorder in a conceptual manner to generate clinical and dynamic signs of NPD that may be used reliably for differential diagnosis.

This review examines the available empirical data as to assessment of NPD and provides both the clinician and researcher with valuable diagnostic information about NPD in a concise manner. Because of the wide range of assessment methods the different types of measures are covered in three sections:

1. Semi-structured interviews which include information on descriptive/phenomenological studies.
2. Self-report inventories.
3. Projective techniques.

Issues of reliability, validity, and clinical utility are examined for each instrument (or scale). An overview of the relative advantages, disadvantages, and empirical support for each method of assessment in the diagnosis of NPD is presented in a discussion after the review of the salient literature.

SEMISTRUCTURED INTERVIEWS

Table 1 summarizes 22 studies that used structured clinical interviews with references to the various diagnostic efficiency statistics, median sensitivity (SN; the ability of the test to correctly identify individuals with the disorder), specificity (SP; the ability of the test to correctly identify nonpersonality disordered individuals as nonpersonality disordered), positive predictive power (PPP), and negative predictive power (NPP; the probability that an individual has or does not have a personality disorder when the test identifies him/her as having or not having a personality disorder).

STRUCTURED INTERVIEW FOR THE DSM PERSONALITY DISORDERS (SIDP)

The SIDP was developed by Pfohl, Stangl, and Zimmerman (1983) and was designed to diagnose *DSM-III* personality disorders. It is composed of 160 questions, grouped into 16 areas of functioning. Questions are organized in this manner to decrease redundancy of a grouping based on specific personality disorders and to provide for a smoother flowing interview. All questions are directed to specific Axis II criteria, and

TABLE 1. Diagnostic Efficiency Statistics for Structured Clinical Interview Measures of the DSM Narcissistic Personality Disorder

Study	Instrument	Sen.	Spec.	PPP	NPP	Kappa NPD	ICC for NPD	Kappa Any PD
Stangl et al., 1985	SIDP	—	—	—	—	.85	—	.71
Pfohl et al., 1986	SIDP	.70	.86	.20	.99	.52	—	—
Dubro et al., 1988	SIDP	—	—	—	—	.68	—	—
Brent et al., 1990	SIDP	—	—	—	—	.32	—	—
Hogg et al., 1990	SIDP	—	—	—	—	—	.77	.59
Miller et al., 1992	SIDP	.24	.88	.29	1.00	—	—	—
Turley et al., 1992	SIDP	—	—	—	—	—	—	.91
Trull et al., 1993	SIDP-R	—	—	—	—	—	.66	.76
Loranger et al., 1987	PDE	—	—	—	—	—	.94	.80
Standage et al., 1988	PDE	—	—	—	—	.63	—	.41
Loranger et al., 1991	PDE	—	—	—	—	.36 ^a	—	.55 ^a
Pilkonis et al., 1991	PDE	—	—	—	—	—	—	.79
Ames-Frankel et al., 1992	PDE	—	—	—	—	.27	—	.48
Skodol et al., 1988	SCID-II	—	—	.75	.38	—	—	—
Stanley et al., 1990	SCID-II	—	—	—	—	.63	—	—
Brooks et al., 1991	SCID-II	—	—	—	—	.78	—	—
Renneberg et al., 1992	SCID-II	—	—	—	—	—	—	.75
Guthrie & Mobley, 1994	SCID-II	—	—	—	—	—	—	.97
First et al., 1995	SCID-II	—	—	—	—	.43 ^a	—	.51 ^a
Widiger et al., 1987	PIQ	—	—	—	—	.77	.82	—
Widiger et al., 1990	PIQ-II	—	—	—	—	.70	—	—
Gunderson et al., 1990	DIN	.63	.88	.68	.85	.52	.88	—

Note. Sen. = Sensitivity; Spec. = Specificity; PPP = Positive Predictive Power; NPP = Negative Predictive Power; Kappa NPD = Interrater agreement for Narcissistic PD; ICC = Interclass correlation for dimensionalized NPD scores; Kappa Any PD = Interrater agreement for presence or absence of any DSM PD; SIDP = Structured Interview for the DSM Personality Disorders; PDE = Personality Disorders Examination; SCID-II = Structured Clinical Interview for DSM-III-R Axis II; PIQ = Personality Interview Questions; DIN = Diagnostic Interview for Narcissism.

^aInterrater Kappas based upon test-retest Methodology.

a list of relevant personality disorder criteria at three levels of severity is listed at the end of each section. Early work on the SIDP (Pfohl, et al. 1983) has demonstrated criterion-based reliability and validity with *DSM-III* Axis II disorders and moderate levels of interrater reliability in determining presence or absence of a personality disorder diagnosis (Dubro, Wetzler, & Kahn, 1988; Stangl, Pfohl, Zimmerman, Bowers, & Corenthal, 1985). However, in both these studies, NPD occurred in a small number of diagnosed cases (3 and 5, respectively).

In a study by Pfohl, Coryell, Zimmerman, and Stangl (1986), overall reliability in the diagnosis of NPD was modest with individual high criteria found to be "grandiose uniqueness" and "interpersonal exploitativeness," while the two least reliably rated criteria were "response to criticism" and "lack of empathy." In addition, the *DSM-III* monothetic criteria were found to have much higher interrater agreement than polythetic criteria for rating presence of *DSM-III* criteria for NPD. Prevalence of NPD based on the SIDP was found to be 16% in a sample made up predominantly of inpatients, whereas Reich and Troughton (1988) found prevalence to be much smaller in a group of outpatients (1.2%).

Miller, Streiner, and Parkinson (1992) evaluated the ability of the SIDP in the identification of personality disorders using 122 psychiatric patients. For this sample, 12 patients scored positive for NPD on the SIDP. Diagnostic efficiency statistics for this measure in the assessment of NPD demonstrated poorer results compared with identification of cluster B personality disorders. This suggests that the SIDP is better suited for identifying those who do not have NPD, than in the identification of those who do have this disorder. However, results of this study also indicate that the SIDP can accurately identify patients diagnosed within the cluster B personality disorders (antisocial, borderline, histrionic, narcissistic).

Two recent studies have examined use of the SIDP in the dual diagnosis of psychiatric inpatients. Hogg, Jackson, Rudd, and Edwards (1990) determined prevalence of personality disorders and personality disorder traits in 40 recent onset schizophrenic patients. The SIDP proved effective in determining presence or absence of personality disorders. Interrater reliabilities for the SIDP narcissism trait ratings were significant; the mean SIDP narcissism trait rating for this sample was 0.32. Turley, Bates, Edwards, and Jackson (1992) also used the SIDP to assess 21 recent onset Bipolar disordered patients for personality disorders. Again, interrater agreement on the SIDP for personality disorder diagnosis was high. These studies suggest that the SIDP can be used to accurately assess narcissistic traits within groups of psychiatric inpatients with dual diagnoses. However, NPD was only diagnosed for 1 patient in the Hogg et al. (1990) study and for 2 patients in the Turley et al. (1992) study. Also, interrater agreement for the NPD diagnosis on the SIDP was found to be the lowest of any personality disorder for a sample of adolescent psychiatric patients (Brent, Zelenak, Bukstein, & Brown, 1990).

Blashfield, Blum, and Pfohl (1992) assessed changes in diagnostic criteria from the *DSM-III* to the *DSM-III-R*, using the SIDP and SIDP-R. In an evaluation of *DSM-III* and *DSM-III-R*, a relatively good Kappa was obtained for the agreement of NPD criteria ($k = 0.571$), the highest of the personality disorders in this study. Although NPD was diagnosed infrequently (6% to 8%, respectively), concordance between the *DSM-III* and *DSM-III-R* was relatively high. Changing from a partial monothetic definition in the *DSM-III* to a polythetic definition in the *DSM-III-R* apparently led to relatively little change in the application of this diagnosis. Likewise, Trull and Goodwin (1993) found that the mean intraclass correlation coefficients comparing independent raters for the 11 personality disorder criteria sets, as measured by the SIDP-R, was 0.76. This indicates very high agreement among the interviewers concerning the number of *DSM-III-R* personality disorder criteria each patient met. Intraclass correlation coefficients of SIDP-R scores for NPD were relatively stable over a 6-month period. The SIDP-R identified 41% of an adult outpatient sample as meeting *DSM-III-R* criteria for at least one personality disorder and 18% as meeting criteria for two or more personality disorders. However, within this group, the number of patients with a diagnosis of NPD was again quite low (4 cases).

In relation to self-report measures of Axis II disorders, the SIDP has been found to identify patients as having NPD less frequently than the Personality Diagnostic Questionnaire (PDQ) would likewise indicate (Zimmerman & Coryell, 1990) but more frequently than the Millon Clinical Multiaxial Inventory (MCMI) and Minnesota Multiphasic Personality Inventory-Personality Disorder Scales (MMPI-PD) in a sample of inpatients (Miller, Streiner, & Parkinson, 1992). However, the SIDP has also been found to diagnose NPD less frequently than the MCMI, using a base rate of > 65 or 75 as the criterion in a sample of outpatients (Torgersen & Alnaes, 1990). Also, the SIDP diagnosed NPD less frequently than the MCMI-II in determining presence of personality disorders

with recent onset bipolar disorder patients (Turley et al., 1992). Dimensional scores of the SIDP for the NPD trait have been found to be significantly correlated with the narcissistic personality disorder scale of the PDQ (PDQ-NPD; Zimmerman & Coryell, 1990), the PDQ-R-NPD scale (Yeung, Lyons, Waternaux, Faraone, & Tsuang, 1993) and with the narcissistic personality disorder scale of the MCMI (MCMI-N; Hogg, Jackson, Rudd, & Edwards, 1990; Torgersen & Alnaes, 1990). It appears that the SIDP/SIDP-R will diagnose NPD less frequently than self-report measures across many different settings. However, it does appear that dimensional scores for narcissism on a number of different self-report measures are significantly related to SIDP/SIDP-R criteria for NPD.

PERSONALITY DISORDER EXAMINATION

The Personality Disorder Examination (PDE; Loranger, 1988) is a structured interview developed for the assessment of Axis II disorders. The PDE arranges 328 items designed to assess *DSM-III-R* criteria in six content areas (Affect, Impulse Control, Interpersonal relations, Reality Testing, Self, and Work). The PDE is structured in this manner to facilitate the flow of a regular clinical interview. All items are rated on a 3-point scale (absent or normal = 0, exaggerated or accentuated = 1, meets criteria for pathological level = 2) and they are tabulated to determine the number of criteria that are met. A dimensional score (representing the amount of the personality trait that is present) and categorical diagnosis are derived. Also, provision is made for assigning a diagnosis as probable (clinical subthreshold) in addition to a definite *DSM-III-R* diagnosis.

Loranger, Susman, Oldham, and Russakoff (1987) found that interrater reliabilities of the dimensional scores for NPD on the PDE were very high ($r = 0.94$), as was the agreement for presence or absence of NPD (97%) using *DSM-III* criteria. However, only 3 cases of NPD were present in this preliminary study. Similarly, Standage and Ladha (1988) found a high level of interrater agreement for NPD on the PDE. However, like the Loranger et al. (1987) study, the sample size of this investigation was small (20 total patients). Using a large sample of inpatients Loranger et al. (1991) found the PDE to be useful in examination of the longitudinal effects of personality disorder criteria. No statistics are given for NPD specifically, but the intraclass correlation coefficients of interrater reliability measuring individual personality disorder criteria at initial interview was 0.93, whereas interrater reliability at the follow-up interview was 0.94. Stability of the personality disorder criteria ratings assigned to the patients had a median correlation of 0.72.

Several studies have investigated the diagnostic agreement of *DSM-III-R* personality disorders of the PDE with other structured interviews and self-report measures. Specifically concerning NPD, one study (Hyler, Skodol, Kellman, Oldham, & Rosnick, 1990) found moderate agreement between the PDE diagnosis of NPD with that of the Structured Clinical Interview for the *DSM-III-R* Axis II (SCID-II) and the PDQ-R. Whereas the SCID-II and PDE exhibited similar prevalence (15 and 19, respectively), the PDQ-R diagnosed more cases (30) of NPD compared with the two structured interviews. A subsequent study (Oldham, Skodol, Kellman, Hyler, Rosnick, & Davies, 1992) again found that the NPD diagnoses utilizing the SCID-II and PDE were highly similar for 100 consecutively admitted inpatients (17% and 20%, respectively). The Ames-Frankel et al. (1992) study found the NPD score on the PDE to be relatively low when comparing evaluations after a 6-week interval for a sample of 30 bulimic outpatients. Whereas Pilkonis, Heape, Ruddy, and Serrao (1991) found that the overall interrater reliability of the PDE was excellent they also observed that the clinical validity of the PDE was poor (SN = 0.71, SP = 0.58, K = 0.28), as determined by examining

agreement between this measure and consensual diagnoses formulated using the LEAD (Longitudinal, Expert, All Data) standard proposed by Spitzer (1983).

Agreement between the PDE and the revision of the PDQ (PDQ-R) is poor for NPD. Hunt and Andrews (1992) found that overall the PDE indicated that only 3 patients met criteria for *DSM-III-R* Axis II diagnosis, where 27 of the 40 patients in this sample met diagnostic criteria for a personality disorder as measured by the PDQ-R. The PDE did not identify any of 40 patients as having NPD, whereas the PDQ-R diagnosed 10 patients with NPD.

Soldz, Budman, Dembry, and Merry (1993) compared the concurrent validity of the MCMI-II and the PDE. Using a sample of 97 outpatients, 66 were assigned personality disorders by the MCMI-II. This was closer to the number of patients assigned a probable or definite diagnosis (75) by the PDE rather than the number given to those patients with just a definite diagnosis alone (43) by the PDE. The MCMI-II diagnosed 11 patients with NPD, whereas the PDE diagnosed only 2 patients as definite NPD. Overall diagnostic agreement between the MCMI-II and PDE about the presence or absence of any personality disorder was significant, as was the relationship between the two instruments in the assessment of cluster B personality disorders. However, agreement between these two instruments in the diagnosis of NPD was very low and not significant. Agreement between the PDE and MCMI-II dimensional scores for NPD was also explored and a significant relationship was found. However, whereas the dimensional scores for the NPD scale on each measure were significantly related to one another, the MCMI-II Schizoid, Histrionic, Passive-Aggressive, Paranoid, Antisocial, and Borderline scales were also significantly related to the dimensional PDE NPD score, with the latter three scales having higher correlations than the NPD scale of the MCMI-II.

STRUCTURED CLINICAL INTERVIEW FOR *DSM-III-R* AXIS II (SCID-II)

The Structured Clinical Interview for *DSM-III-R* Axis II personality disorders (SCID-II) developed by Spitzer, Williams, and Gibbon (1987), assesses each *DSM-III-R* personality disorder in succession. One or two questions are provided that directly assess each personality disorder symptom/criterion. Use of the SCID-II to diagnose NPD resulted in the lowest overall diagnostic power (0.45), compared with its use for the other 10 personality disorders when evaluated with a LEAD approach to diagnosis (Skodol, Rosnick, Kellman, Oldham, & Hyler, 1988). This result was very similar to the work of First and colleagues (1995), who found that the short-interval test-retest reliability of the SCID-II in the diagnosis of NPD was relatively low (0.43). However, other researchers have found the NPD scale of the SCID-II to have good interrater reliability for outpatients being treated for Axis I disorders (Brooks et al., 1991; Guthrie & Mobley, 1994; Stanley, Turner, & Borden, 1990).

In comparison to other assessment instruments, prevalence of NPD utilizing the SCID-II and PDE was highly similar for 100 consecutively admitted inpatients (17% and 20%, respectively; Oldham et al., 1992). In a comparison of the SCID-II with the PDQ-R and the PDE in the diagnosis of *DSM-III-R* personality disorders in 87 patients (Hyler et al., 1990), the SCID-II yielded more diagnoses, on average than the PDE but less than the PDQ-R. Specifically concerning NPD, the PDQ-R diagnosed 30 patients, whereas the SCID-II identified 15 and the PDE 19. Agreement between the PDQ-R and the SCID-II in the diagnosis of NPD was significant as was the relationship between the SCID-II and PDE. In contrast to the previous study, Hills (1995) and Guthrie and Mobley (1994) found poor agreement between the SCID-II with either the MMPI-2 or MCMI-II scales for the diagnosis of NPD.

Renneberg, Chambless, Dowdall, Fauerbach, and Gracely (1992) examined the concurrent validity of personality disorder diagnosis, comparing the SCID-II and MCMI-II. The interrater reliability coefficients for presence or absence of any personality disorder with the SCID-II was 0.75. For the individual personality disorder categories, reliability coefficients ranged from 0.61 to 0.81. Kappa coefficients were not computed for NPD because this disorder was not diagnosed 3 or more times in this sample. Despite satisfactory agreement between the MCMI-II and the SCID-II, present data show very limited support for the concurrent validity of these two measures of personality disorders. Although the Kappa coefficients between the measures were generally positive and significant, they were quite low for validity coefficients, with few exceptions. However, it should be noted that employing a BR of > 84 as the MCMI-II criterion led to overall prevalence rates similar to that of the SCID-II (56% and 57%, respectively), but with no better agreement concerning which patients have a specific personality disorder.

CLINICAL DIAGNOSIS UTILIZING *DSM-III/DSM-III-R* CRITERIA

Some investigators have successfully assessed DSM diagnostic criteria during a clinical interview. Widiger, Trull, Hurt, Clarkin, and Frances (1987) utilized a semistructured interview (PIQ) based on the diagnostic criteria for the 11 *DSM-III* personality disorders. Interrater reliability for interviewers' ratings of NPD was very high ($K = 0.77$). In a later revision adapted for the *DSM-III-R*, Widiger, Freiman, & Bailey (1990) also found high interrater reliability for the diagnosis of NPD. Interrater reliability for number of symptoms recorded for NPD patients was 0.82. Vaillant and Drake (1985) also employed a similar method to assess patients for the presence or absence of a personality disorder with significant results as well ($K = 0.77$). This suggested that integrating a number of criteria for a specific or selected number of personality disorders into a clinical interview can yield diagnoses in a highly reliable manner.

RETROSPECTIVE *DSM-III/DSM-III-R* DIAGNOSIS UTILIZING CHART INFORMATION

Several studies have utilized chart information to retrospectively rate *DSM-III/DSM-III-R* diagnostic criteria for NPD. In a review of issues and research methods for diagnosing personality disorders, Zimmerman recently stated "Patients whose conditions are diagnosed according to retrospective chart review are likely to be prototypic examples of the PD [personality disorder]" (1994, p. 232). Presence or absence of symptoms is determined in a retrospective review of patient records by raters who are blind to patient diagnosis. Each rater then identifies presence or absence of each *DSM-III* symptom for a specific personality disorder. Employing this method, Morey (1985), using a point-biserial correlation that was corrected for the number of symptoms, found a moderate reliability coefficient of 0.58 across each personality disorder. However, diagnostic criteria for NPD appeared to have a much higher internal consistency than average reliability across all the personality disorders. All the NPD symptoms had highly significant associations with the NPD total scale score, and only one symptom, "idealized/devalued relationships," was correlated significantly with the BPD (0.33) and manic (0.35) total scale scores. Also, BPD symptoms of "unstable and intense relationships," as well as "affective instability," were significantly related to the NPD total scale score (both at 0.40). Finally, a number of schizophrenic/schizotypal criteria and symptoms were found to be negatively associated with NPD.

Plakun (1987) also utilized retrospective *DSM-III* diagnoses, based on portions of case records, made blind to patient identity and diagnosis. Interrater reliability of

independent ratings of 25 randomly selected charts of hospitalized patients indicated complete rater agreement on the presence or absence of NPD. Interrater agreement for presence or absence of individual NPD criteria ranged between 75% for "preoccupation with fantasies of success" and 90% for "cool indifference toward others."

Other authors have also successfully used this method in research with Axis II disorders (Stone, 1989). Using an independent retrospective *DSM-III* diagnosis by two raters who were blind to chart diagnosis, Plakun (1989) found high reliability coefficients for presence or absence of Axis II personality disorders. A similar method was employed by Castlebury et al. (in press) where classification of *DSM-IV* Cluster B personality disorders was based on an extensive review of chart materials. Interrater agreement for each of the four Cluster B personality disorders in this study was found to be quite high (≥ 0.80). Clinical chart abstracts have also been used to rate BPD patients on presence or absence of *DSM-III-R* criteria for NPD. In this study, interrater reliabilities for *DSM-III* BPD criteria were high. However, Kappa was not calculated for NPD ratings (McGlashan & Heinssen, 1989).

DIAGNOSTIC INTERVIEW FOR NARCISSISM

The *Diagnostic Interview for Narcissism* (DIN) is the first instrument that has been developed from and applied to clinically diagnosed narcissistic patients. The DIN (Gunderson & Ronningstam, 1987) was developed from a review of salient clinical literature (Ronningstam, 1988), clinical experience, and a systematic examination of NPD patients (Ronningstam & Gunderson, 1988). These authors developed a semi-structured interview consisting of 33 statements in five sections: grandiosity (I), interpersonal relations (II), reactivity (III), affect and mood states (IV), and social and moral adaptation (V). Interrater reliability coefficients for each section and for the total interview score were very high. For a sample of 82 patients (24 NPD, 54 non-NPD), a correlation of the 33 statements with the total score revealed good internal consistency (0.81). Correlation of the statements with their corresponding section scores also revealed good consistency for Sections I (0.76), II (0.60), and IV (0.65). Lower correlations were found for Sections III (0.44), and V (0.49). Individual correlation coefficients for each of the Sections with the total interview score revealed high consistency (0.66). Using a total scaled score of 9 as a cutoff point, the DIN separated the NPD from non-NPD patients with a high degree of accuracy (Gunderson et al., 1990).

Further evaluation of the DIN mean scores of patients with or without NPD showed significant differences on 7 of the 8 statements in Section I, 3 of the 9 statements in Section II, 1 of the 5 statements in Section III, and 3 of the 6 statements in Section V. None of the mean scores for the 5 statements in Section IV exhibited significant differences between the two groups (Ronningstam & Gunderson, 1990). In addition, these authors further examined whether the DIN would be able to differentiate NPD from the related BPD (Ronningstam & Gunderson, 1991). Using the DIN in the differential diagnosis of NPD and BPD patients, a similar pattern of results was obtained with those of the previous study and show the mean scores on several of the 33 statements comprising the DIN to be significantly different between the two groups, with Section I (grandiosity) the most robust in discriminating between NPD and BPD.

SELF-REPORT MEASURES

Table 2 and Table 3 summarize the concurrent validity data for the PDQ, PDQ-R, MMPI, MMPI-2, MCMI, and MCMI-II, along with NPD internal consistency data and

TABLE 2. Concurrent Validity and Psychometric Properties of Self-Report Measures of the DSM Narcissistic Personality Disorder: Structured Interviews and Clinician Diagnosis as Criterion Measures

Study	Instrument	NPD Internal Consistency	Test-Retest r	Concurrent Validity r (Criterion Measure)
Hyer et al., 1989	PDQ-NPD	.57	—	.31 (Clinician DSM-III Dx)
Zimmermann et al., 1990	PDQ-NPD	—	—	.26 (SIDP-NPD)
Hyer et al., 1990	PDQ-R-NPD	—	—	.34 ^a (SCID-II-NPD)
Hyer et al., 1990	PDQ-R-NPD	—	—	.42 ^a (PDE-NPD)
Dowson, 1992	PDQ-R-NPD	—	—	.42 (DSM-III-R NPD ratings)
Yeung et al., 1993	PDQ-R-NPD	—	—	.15 (SIDP-NPD)
Trull, 1993	PDQ-R-NPD	.49	.66	—
Trull et al., 1993	PDQ-R-NPD	—	.60	—
Castlebury et al., in press	MMPI-2-NPD-O	—	—	.55 (DSM-IV NPD criteria)
Castlebury et al., in press	MMPI-2-NPD-NO	—	—	.47 (DSM-IV NPD criteria)
Hills, 1995	MMPI-2-NPD-NO	—	—	-.01 (SCID-II-NPD)
Hogg et al., 1990	MCMI-NPD	—	—	.34 (SIDP-NPD)
Torgersen et al., 1990	MCMI-NPD	—	—	.18 (SIDP-NPD)
Chick et al., 1993	MCMI-NPD	—	—	.09 (DSM-III-R NPD criteria)
Soldz et al., 1993	MCMI-II-NPD	—	—	.41 (PDE-NPD)
Hills, 1995	MCMI-II-NPD	—	—	.11 (SCID-II-NPD)

Note. Internal consistency = Coefficient Alpha; Test-retest intervals \leq 6 months; Concurrent validity = correlation coefficient; PDQ = Personality Disorder Questionnaire; PDQ-R = Personality Disorder Questionnaire-Revised; MMPI-NPD-O = Minnesota Multiphasic Personality Inventory Narcissistic PD Overlapping Scale; MMPI-NPD-NO = Minnesota Multiphasic Personality Inventory Narcissistic PD Non-Overlapping Scale; MCMI-NPD = Millon Clinical Multiaxial Inventory, Narcissistic PD Scale; MCMI-II-NPD = Millon Clinical Multiaxial Inventory-II, Narcissistic PD Scale; SIDP-N = Structured Interview for the DSM Personality Disorders; PDE-N = Personality Disorders Examination-Narcissistic PD scale; SCID-II-NPD = Structured Clinical Interview for DSM-III-R Axis II-Narcissistic scale.

^aConcurrent validity = a Kappa coefficient.

test-retest reliability data when available. Table 4 summarizes the diagnostic efficiency statistics for these same measures.

PERSONALITY DIAGNOSTIC QUESTIONNAIRE (PDQ/PDQ-R)

The PDQ, developed by Hyler, Rieder, Spitzer, and Williams (1982), is a 163-item true-false inventory that typically takes less than 1 hour to complete. The items of the PDQ are specifically designed to assess the diagnostic criteria for *DSM-III* personality disorders. Each personality disorder criterion is represented by one or more items on the PDQ. The PDQ provides a score for each of the *DSM-III* personality disorders representing a number of criteria met for each disorder and a total PDQ summation score of all pathological responses to PDQ items, representing a measure of overall personality disturbance.

The total PDQ score can be used as an indicator of overall personality disturbance, although its ability to distinguish between specific personality disorders may be less

TABLE 3. Psychometric Properties of Self-Report Measures of the DSM Narcissistic Personality Disorder: Self Report Instruments as Criterion Measures

Study	Instrument	NPD Internal Consistency	Test-Retest r	Concurrent Validity r (Criterion Measure)
Morey et al., 1985	MMPI-NPD-O	.77K	—	—
Morey et al., 1985	MMPI-NPD-NO	.71K	—	—
Morey et al., 1988	MMPI-NPD-O	—	—	.68 (MCMI-NPD)
Striener et al., 1988	MMPI-NPD-O	—	—	.66 (MCMI-NPD)
Dubro et al., 1989	MMPI-NPD-O	—	—	.55 (MCMI-NPD)
McCann, 1989	MMPI-NPD-O	—	—	.78 (MCMI-NPD)
McCann, 1989	MMPI-NPD-NO	—	—	.64 (MCMI-NPD)
Hurt et al., 1990	MMPI-NPD-O	—	.73	—
McCann, 1991	MMPI-NPD-O	—	—	.65 (MCMI-II-NPD-O)
McCann, 1991	MMPI-NPD-NO	—	—	.50 (MCMI-II-NPD-NO)
Chatham et al., 1993	MMPI-NPD-O	—	—	.66 (MCMI-NPD)
Chatham et al., 1993	MMPI-NPD-O	—	—	.68 (NPI)
Trull et al., 1993	MMPI-NPD-O	—	.71	—
Trull, 1993	MMPI-NPD	.70	.74	—
Schuler et al., 1994	MMPI-NPD-O	—	—	.73 (MCMI-NPD)
Schuler et al., 1994	MMPI-NPD-NO	—	—	.61 (MCMI-NPD)
Hills, 1995	MMPI-2-NPD-NO	—	—	.25 ^a (MCMI-II-NPD)
Wise, 1996	MMPI-2-NPD-O	—	—	.68 (MCMI-II-NPD-O)
Wise, 1996	MMPI-2-NPD-NO	—	—	.68 (MCMI-II-NPD-O)
Prifitera et al., 1984	MCMI-NPD	—	—	.66 (NPI)
Piersma et al., 1986	MCMI-NPD	—	.61	—
Chatham et al., 1993	MCMI-NPD	—	—	.75 (NPI)

Note. Internal consistency = Coefficient Alpha or K = KR-20; Test-retest intervals \leq 6 months; Concurrent validity = correlation coefficient; PDQ = Personality Disorder Questionnaire; PDQ-R = Personality Disorder Questionnaire-Revised; MMPI-NPD-O = Minnesota Multiphasic Personality Inventory Narcissistic PD Overlapping Scale; MMPI-NPD-NO = Minnesota Multiphasic Personality Inventory Narcissistic PD Non-Overlapping Scale; MCMI-NPD = Millon Clinical Multiaxial Inventory, Narcissistic PD Scale; MCMI-II-NPD = Millon Clinical Multiaxial Inventory-II, Narcissistic PD Scale; SIDP-N = Structured Interview for the DSM Personality Disorders; PDE-N = Personality Disorders Examination-Narcissistic PD scale; SCID-II-N = Structured Clinical Interview for DSM-III-R Axis II-Narcissistic scale.

^aConcurrent validity = a Kappa coefficient.

effective (Hyler et al., 1988). In one sample of 552 patients, the relationship between clinicians' diagnoses of personality disorder and the PDQ showed a general lack of agreement between clinical and PDQ diagnoses of *DSM-III* personality disorders. It was observed that the PDQ generated far more diagnoses than were based on clinical judgment. Clinicians made 0.7 personality diagnoses per patient; the PDQ generated 2.4 diagnoses per patient for the entire 552 patient sample. Clinicians reported 12% of the sample was found to have personality disorders, and for this group, the clinicians made 1.2 diagnoses per patient. As for this same 12% of patients, the PDQ made an average of 3.0 personality diagnoses. However, NPD was the only diagnosis made more frequently by clinicians than by the PDQ, with the clinicians indicating the diagnosis three times more frequently than the PDQ (33 and 11, respectively). Also, the correlation between the clinicians' scaled ratings of NPD criteria and the PDQ dimensional scores for narcissism were significantly related to one another (Hyler et al., 1989).

TABLE 4. Diagnostic Efficiency Statistics for Self-Report Measures of the DSM Narcissistic Personality Disorder

Study	Instrument	Sen.	Spec.	PPP	NPP	Kappa NPD	Criterion Measure
Dubro, 1988	PDQ-NPD	.59	.96	—	—	—	DSM-III-R NPD criteria
Hyer et al., 1989	PDQ-NPD	—	—	.09	.98	.10	DSM-III NPD criteria
Zimmermann et al., 1990	PDQ-NPD	—	—	—	—	.00	SIDP-NPD
Hyer et al., 1990	PDQ-R-NPD	.89	.71	.27	.98	.30	SCID-II-NPD and PDE-NPD
Hyer et al., 1990	PDQ-R-NPD ^a	.68	.79	.57	.86	.44	SCID-II-NPD or PDE-NPD
Yeung et al., 1993	PDQ-R-NPD	.00	.95	.00	.97	-.04	SIDP-NPD
Guthrie & Mobley, 1994	PDQ-R-NPD	.75	.75	.19	.97	—	SCID-II-NPD
Miller et al., 1992	MMPI-NPD-O	.33	.99	.87	.94	—	SIDP-NPD
Guthrie & Mobley, 1994	MMPI-NPD-O	.00	.98	.00	.93	—	SCID-II-NPD
Hills, 1995	MMPI-2-NPD-NO	.00	.96	.00	.93	-.05	SCID-II-NPD
Castlebury et al., in press	MMPI-2-NPD-O	.46	.60	.11	.91	.02	DSM-IV NPD criteria
Castlebury et al., in press	MMPI-2-NPD-NO	.39	.66	.11	.91	.02	DSM-IV NPD criteria
Hogg et al., 1990	MCMF-NPD	—	—	—	—	.13	SIDP-NPD
Torgersen et al., 1990	MCMF-NPD	.15	.97	.22	.96	—	SIDP-NPD
Miller et al., 1992	MCMF-NPD	1.00	1.00	1.00	1.00	—	SIDP-NPD
Chick et al., 1993	MCMF-NPD	.50	.80	.05	.99	—	DSM-III-R NPD criteria
McCann, 1990	MCMF-II-NPD	—	.78	—	—	—	SIDP-NPD
Turley et al., 1992	MCMF-II-NPD	—	—	—	—	.02	SIDP-NPD
Soldz et al., 1993	MCMF-II-NPD	.00	.88	.00	.98	-.04	PDE Definite NPD
Soldz et al., 1993	MCMF-II-NPD	.18	.90	.18	.90	.08	PDE Probable NPD
Guthrie & Mobley, 1994	MCMF-II-NPD	.50	.90	.29	.96	—	SCID-II-NPD
Hills, 1995	MCMF-II-NPD	.38	.84	.14	.95	.11	SCID-II-NPD

Note. Sen. = Sensitivity; Spec. = Specificity; PPP = Positive Predictive Power; NPP = Negative Predictive Power; Kappa NPD = Interrater agreement for Narcissistic PD; ICC = Interclass correlation for dimensionalized NPD scores; PDQ-R = Personality Disorder Questionnaire-Revised; MMPI-2-N-NO = Minnesota Multiphasic Personality Inventory Narcissistic-2 NPD Non-Overlapping Scale; MCMF-II-NPD = Millon Clinical Multiaxial Inventory-II, Narcissistic PD Scale; SIDP-NPD = Structured Interview for the DSM Personality Disorders NPD scale; PDE-N = Personality Disorders Examination NPD scale; SCID-II-N = Structured Clinical Interview for DSM-III-R Axis II-NPD scale.
^aPatients with probable NPD on either SCID-II or PDE.

Zimmerman and Coryell (1990) examined comparability of the PDQ and SIDP in a sample of 697 relatives of psychiatric patients and healthy controls. Significantly more individuals had a personality disorder according to the SIDP than the PDQ (94 vs. 72, $p < 0.05$). However, multiple personality disorders were more frequently diagnosed on the PDQ (55 vs. 29, $p < 0.001$). Within this sample, the PDQ diagnosed 3 cases of NPD, whereas the SIDP identified none of the cases as having this disorder. Although the average dimensional scores on the two measures for each corresponding personality disorder were generally high, the correlation of dimensional scores for NPD was the second lowest, but was still significant. The PDQ dimensional score was more than three times higher than the SIDP score (0.5 vs. 1.6). However, concordance for categorical diagnoses was poor because the SIDP failed to diagnose any cases of NPD. Yeung et al. (1993) found similar results when evaluating the diagnostic agreement of the PDQ-R with the SIDP. These authors also report a small, significant, correlation between the dimensional scores of these two measures but, again, the categorical agreement for these measures was quite poor.

Dubro et al. (1988) also evaluated the diagnostic efficiency of the PDQ in relation to SIDP diagnosis when used as screening instruments to identify presence or absence of a personality disorder. In this role, the PDQ had excellent sensitivity, although it tended to overdiagnose the presence of personality disorders. The PDQ diagnosis correctly predicted the criterion diagnosis 63% of the time and correctly predicted absence of a personality disorder with 93% accuracy. With the exception of specificity, the diagnostic efficiency statistics were not calculated for NPD because there was a small number of diagnosed cases in this sample ($n = 3$). Specificity rate was calculated for NPD and this was found to be high (0.78). Findings for this study seemed to suggest that, as a screening instrument for the detection of a personality disorder, the PDQ is an effective measure.

Items of the PDQ were revised to better assess *DSM-III-R* criteria (Hyer & Rieder, 1987). The agreement between the PDE and its more recent revision, the PDQ-R, has been shown to be poor for NPD. As reported earlier, Hunt and Andrews (1992) found that the overall diagnostic agreement for NPD between the PDE and PDQ-R was poor. The PDE did not identify any of the 40 patients in this study as having NPD, whereas the PDQ-R diagnosed 10 patients with NPD. A second study (Dowson, 1992) utilized the PDQ-R with 60 psychiatric patients and an informant who was usually a spouse or first-degree relative. Pearson correlation coefficients between the number of *DSM-III-R* criteria, from the PDQ-R, for NPD reported by the patient with the PDQ-R ratings of NPD by informants was 0.42. Agreement of patient and informant ratings of the 9 individual criteria for NPD produced three criteria with significant Kappas. These were Criterion 1 (reacts to criticism with feelings of rage, shame or humiliation), Criterion 6 (has a sense of entitlement), and Criterion 7 (requires constant attention and admiration). Of these three criteria, a "sense of entitlement" appeared to be the aspect of narcissism that was most reliable in using this questionnaire method. Identification of a sense of entitlement by the patient may be a relatively reliable and valid indicator of narcissism. However, on patient self report, NPD criteria were significantly associated with histrionic (0.39), borderline (0.33), and passive-aggressive (0.28) personality disorders. Also, a diagnosis of NPD from informant scores was significantly correlated (0.52), with a comorbid diagnosis of HPD as well.

Test-retest reliability of the PDQ-R was evaluated with 51 adult psychiatric outpatients. The test-retest coefficient for the PDQ-R-N was 0.66 and evaluations of internal consistency were moderate and quite similar at both the index (0.49) and at 3-month follow-up (0.50) evaluations (Trull, 1993; see Table 2). Intraclass correlation coefficients

representing stability of the narcissism subscale over a 6-month period was also high (Trull & Goodwin, 1993).

Hylar et al. (1990) compared the PDQ-R with the SCID-II and PDE in the diagnosis of *DSM-III-R* personality disorders in 87 patients. The PDQ-R diagnosed more patients as having each of the personality disorders than did either of the structured interviews. Specifically concerning NPD, the PDQ-R diagnosed 30 patients with NPD; the SCID-II identified 15 and the PDE identified 19. Agreement between the PDQ-R and the SCID-II in the diagnosis of NPD and between the PDQ-R and PDE were all significant. The Kappa, SP, SN, PPP, and NPP of the PDQ-R and the diagnosis of NPD were computed under two different conditions. The first condition is when a patient was diagnosed according to both the SCID-II and the PDE as having NPD (labeled by the authors as definite NPD); the second condition is when a patient was diagnosed with NPD according to either the SCID-II or the PDE as having NPD (labeled by the authors as probable NPD). The diagnostic efficiency statistics are seen in Table 4. In the definite NPD condition, 9 patients were identified, and in the probable NPD condition, 25 patients were identified.

MINNESOTA MULTIPHASIC PERSONALITY INVENTORY-PERSONALITY DISORDER SCALES (MMPI-PD)

The MMPI-PD scales were developed through a combination of rational and empirical strategies from the original MMPI item pool (Morey et al., 1985). Internal consistency values were found to be high (NPD = 0.77). All 11 *DSM-III* personality disorders have both overlapping scales (O; where items are shared with other personality disorder scales) and nonoverlapping scales (NO; where items are only found on an individual scale). Recently, contemporary norms for adults and adolescents have been reported for both MMPI-PD-O and MMPI-PD-NO scales, in addition to a list of item numbers for the MMPI and MMPI-2 (Colligan et al., 1994).

One of the first studies to assess the diagnostic efficiency of the MMPI-PD scales was conducted in relation to SIDP diagnosis by Dubro et al. (1988). Similar to findings with the PDQ-R, it was found that the MMPI-PD scales can also be a useful screening instrument to identify the presence or absence of a personality disorder. The MMPI-PD scales demonstrated high specificity; predictive accuracy of the presence or absence of a personality disorder was 90%, although the scales overdiagnosed personality disorders. For cluster B diagnoses, the ability of the MMPI-PD scales to correctly identify nonpersonality disordered individuals as not having a personality disorder was fair, but ability of these scales to correctly identify individuals with specific personality disorders was low. Diagnostic efficiency statistics, with the exception of specificity, were not calculated for NPD because there was a small number of diagnosed cases in this sample ($n = 3$). However, specificity rates were calculated for NPD and this statistic was high. A second study by these authors (Dubro & Wetzler, 1989) examined the diagnostic efficiency of the MMPI-PD scales in relation to SIDP diagnosis and MCMI base rate scores. Mean MMPI-NPD scores of personality disordered patients and nonpersonality disordered patients were not significantly different between the two groups. In addition, a comparison of MMPI-NPD scores between cluster B personality disorders compared with other personality disorders from cluster A and C were not significantly different. However, the correlation between the MMPI-NPD and MCMI scale 5 (Narcissism) was significant. Miller, Streiner, & Parkinson (1992) also, estimated the ability of the MMPI-PD scales to identify personality disorders using 122 psychiatric patients (see Table 4). A personality disorder was considered present if a patient's

score was equal to or exceeded a *T* score of 70 on the MMPI-PD scales. Only 1 patient scored positive for NPD on the MMPI-NPD scale.

Chatham and colleagues (1993) compared both narcissistic and nonnarcissistic groups of psychiatric patients formed on the basis of criteria comprised of clinical judgment and scores on the Narcissistic Personality Inventory (NPI; Raskin & Hall, 1979) with the MMPI and MCMI. Significant profile differences between the two groups were found to exist; narcissistic patients had significantly higher mean scores on the NPD scales of both these measures. The MMPI-NPD scale was found to be the most robust of five MMPI subscales designed to assess narcissism. The narcissistic group also produced significantly higher scores on the Pd and Ma MMPI scales, as well as significantly lower scores on the Si scale. The MMPI-NPD scale was significantly correlated with both the NPI and MCMI-N.

Castlebury et al. (1996) explored the diagnostic utility of the MMPI-2-NPD scales with a sample of 53 outpatients diagnosed with a Cluster B personality disorder (13 NPD) contrasted with a group of non-Cluster B personality disorders (20) and a non-clinical population (67). Scores for both the overlapping and nonoverlapping scales of the MMPI-2-NPD were used in calculating diagnostic efficiency statistics. Whereas both versions of the MMPI-2-NPD scales showed good convergent validity with *DSM-IV* NPD criteria, these scales ranged from fair to poor in the ability to correctly classify NPD depending on which comparison group was utilized.

Several studies have also examined the convergent validity of the MMPI-NPD and MCMI-N scales. All found a moderate to high level of convergent validity and significant correlation coefficients (Hills, 1995; Morey & Le Vine, 1988; Streiner & Miller, 1988; see Table 3). Also, scores from both the overlapping and nonoverlapping MMPI-NPD scales have also been compared with the MCMI-N. The correlation between MMPI-NPD-O and MMPI-NPD-NO with MCMI-N base rate scores were very high (McCann, 1989). This relationship has also been noted to exist between overlapping and nonoverlapping scales of the MMPI-NPD and the MCMI-II-N items which overlap with other personality scales (MCMI-II-N-O) and those which have no overlap on other scales (MCMI-II-N-NO; McCann, 1991; See Table 3). Correlations between both the overlapping scales of the MMPI-NPD and the MCMI-II-N were higher than the correlation between the nonoverlapping scales of the MMPI-NPD with MCMI-II-N, which were moderate ($r = 0.65$ and $r = 0.50$, respectively). A further study (Schuler, Snibbe, & Buckwalter, 1994) found that both the overlapping and nonoverlapping MMPI-NPD scales were highly correlated with MCMI-N ($r = 0.73$, and $r = 0.61$, respectively). In addition, this study found that the MMPI-PD-O scales were the most accurate in predicting membership in the Cluster B group of personality disorders. These results suggest that both sets of MMPI-NPD scales have similar levels of convergent validity with the respective MCMI-II scale and that the overlapping scales from both measures are related to a higher degree than the nonoverlapping scales from the same measures.

Test-retest reliability and convergent validity of the MMPI-PD scales were evaluated with a sample of 51 adult psychiatric outpatients (Trull, 1993). Internal reliability coefficients for the MMPI-PD subscales exceeded those of the PDQ-R in almost every case at an index evaluation and again after a 3-month follow-up, indicating a greater homogeneity for these scales. In most cases, the MMPI-PD test-retest coefficients were also higher than those of the PDQ-R. The replication of the Hurt, Clarkin, and Morey (1990) reliability values for the MMPI-PD scales are encouraging especially considering that the intervening time period in that study was 3 weeks. As for the MMPI-NPD scale, the reliability coefficients at index evaluation were 0.70, and 0.63 after a 3-month

follow-up. Test-retest correlation of this measure was 0.74. An extremely low and non-significant relationship between the MMPI-NPD scale and the PDQ-R-N scale was also noted. Lack of convergence between the MMPI-NPD scale and the PDQ-R-N may be due to the latter measure being scored directly according to *DSM-III-R* criteria, whereas the MMPI-NPD scale was derived from items based on the *DSM-III*. In a further study with 44 outpatients concerning stability of the MMPI-NPD scale over 6 months, an intraclass correlation coefficient of 0.71 was obtained, representing the stability of the narcissism subscale (Trull & Goodwin, 1993).

The MMPI-NPD scales displayed good ability in identifying those who did not have a personality disorder. However, these scales overdiagnosed NPD which make the results of these scales in the discrimination of different patient populations equivocal. The MMPI-NPD scales were unable to differentiate personality disordered patients from nonpersonality disordered patients but were able to discriminate narcissistic patients from those found to be suffering from nonnarcissistic psychopathology. Several studies have found the convergent validity of the MMPI-NPD and MCMI-N scales to be significantly related. However, the relationship between the MMPI-NPD scale and PDQ-R-N scale was not found to be significant.

MILLON CLINICAL MULTIAXIAL INVENTORY (MCMI/MCMI-II)

The MCMI is a 175-item, true-false inventory of psychopathology (Millon, 1983). This self-report measure is designed to assess both Axis I (9 clinical syndrome scales) and Axis II (11 personality disorder scales) disorders. A positive diagnosis is made when a patient has a Base Rate (BR) score of > 85 on a given index. Piersma (1986) investigated the stability of the personality and symptom scales of the MCMI for a sample of psychiatric inpatients ($N = 151$). Patients were administered the MCMI shortly following admission and shortly preceding discharge ($\bar{X} = 30.4$ days). Results indicated that stability estimates were greater for the personality scales than for the symptom scales. Test-retest stability coefficient of the MCMI-N was 0.61 for this sample.

Chick, Sheaffer, Goggin, and Sison (1993) examined the relationship between elevations on the personality scales of the MCMI and clinician generated *DSM-III-R* diagnoses for 101 psychiatric patients. Personality disorder diagnoses were made by employing a personality symptom checklist that consisted of all the verbatim criteria for personality disorders contained in the *DSM-III-R*. Clinicians who completed checklists were required to have had at least 5 hours of direct contact with the patients who completed the MCMI. Seventy-three patients met the criterion for a personality disorder on the MCMI (base rate > 84), whereas only 46 met the criteria for a personality disorder as defined by the symptom checklist based exclusively on *DSM-III-R* criteria. The sample had a mean base rate score of 59.2 on MCMI-N and 21 subjects exhibited a base rate of greater than 84. However, only 2 of these patients met NPD criteria on the clinician-generated symptom checklist. An examination of the diagnostic efficiency statistics across all of the MCMI personality disorder scales revealed overall low sensitivity, specificity, positive predictive power, negative predictive power, and categorical agreement suggesting that the MCMI may have only limited utility in identifying specific personality disorders. A Spearman correlation coefficient between MCMI-N scores and checklist ratings for NPD criteria was 0.09. These figures suggest that the MCMI tends to overdiagnose Narcissistic Personality Disorder. Furthermore, the authors of this study interpret these results as suggesting that the MCMI personality disorder scales are generally poor in identifying patients who meet *DSM-III-R* criteria for corresponding disorders.

Wetzler and Dubro (1990) also examined diagnostic efficiency of the MCMI with regard to the diagnosis of *DSM-III* axis II personality disorders in comparison with criterion diagnosis by attending psychiatrists. However, this study found diagnostic efficiency statistics of the MCMI for presence or absence of any personality disorder and for diagnosis by axis II cluster to be somewhat higher than assessment of a specific individual personality disorder such as NPD.

Diagnostic efficiency of the MCMI has been evaluated in relation to SIDP diagnosis and the results suggest that the MCMI exhibited excellent sensitivity when used as a screening instrument to identify presence or absence of a personality disorder. For identification of cluster B diagnoses the specificity and sensitivity rates for the MCMI were fair. Diagnostic efficiency statistics with the exception of SP were not calculated for NPD because there were a small number of diagnosed cases in this sample ($n = 3$). Specificity rates were calculated for NPD and were found to be high (Dubro et al., 1988). A second study (Torgersen & Alnaes, 1990) comparing MCMI diagnosis with that of the SIDP for 272 psychiatric outpatients, 81% of whom met criteria for a personality disorder, showed mixed results. Diagnostic efficiency statistics based on SIDP-N for the MCMI-N, using a base rate of 85 for the MCMI-N, do not indicate strong agreement between the SIDP and the MCMI-N for the categorical diagnosis of NPD. However, the MCMI-N score was significantly correlated with SIDP-N.

Hogg and colleagues (1990) studied prevalence of personality disorders and personality disorder traits in 40 recent onset schizophrenic patients to establish degree of concordance between the SIDP and MCMI. As in the previous study, these authors also found the correlation between the NPD trait ratings of the SIDP and MCMI-N base rate scores to be significant. However, they also found that the MCMI diagnosed more of the sample as having a personality disorder compared with the SIDP (Table 4). Miller and colleagues (1992) also assessed the ability of the MCMI to identify personality disorders using 122 psychiatric patients and found 10 patients scoring positive for personality disorders on the MCMI for NPD.

Chatham et al. (1993) found significant profile differences between a group of narcissistic and nonnarcissistic psychiatric patients ($n = 35$ in each group) on the basis of criteria comprised of clinical judgment and scores on the NPI (>25 or <17 , respectively). Comparison of the narcissistic and nonnarcissistic groups on the MCMI revealed that the narcissistic group mean on scale 5 (narcissism) was significantly greater than that of the nonnarcissistic group. However, the narcissistic group scored significantly higher on scales 4 (Histrionic), 6 (Antisocial), P (Paranoid), N (Hypomanic), and T (Drug Abuse), as well as scoring significantly lower on scales 1 (Schizoid), 2 (Avoidant), 3 (Dependent), and S (Schizotypal). The MCMI-N scores were also significantly correlated with both the NPI and MMPI-NPD (see Table 3).

As shown in Table 3, and reported earlier, several studies that assessed the correlation between the MMPI-NPD scales and MCMI-N found these two scales to be highly correlated (Dubro & Wetzler, 1989; Morey & Le Vine, 1988; Streiner & Miller, 1988). Scores from both the overlapping and nonoverlapping MMPI-NPD scales have also been compared with the MCMI-N. Correlations between the overlapping and nonoverlapping MMPI-NPD scales with MCMI-N base rate scores were also found to be very high (McCann, 1989; Schuler, Snibbe, Buckwalter, 1994; see Table 3). Finally, much like the MCMI, correlations between the overlapping and nonoverlapping versions of both the MMPI-NPD scales with MCMI-II-N base rate score were high (McCann, 1991; see Table 3). In addition to the MMPI-NPD, the MCMI-N has also been found to be significantly related to the NPI (Prifitera & Ryan, 1984; see Table 3).

The MCMI-II (Millon, 1987) represents an attempt to address the psychometric shortcomings of the original measure and to make the inventory more comparable to *DSM-III-R* diagnoses. McCann (1990) was one of the first researchers to estimate the diagnostic efficiency of the personality disorder scales of the MCMI-II. MCMI-II scale elevations were compared with clinical diagnoses of *DSM-III-R* personality disorders based on clinicians' primary and secondary Axis II diagnoses as a criterion. Utilizing both of these possible criteria in the diagnosis of NPD, the MCMI-II classification rate had a sensitivity of 0.74 and a specificity of 0.94. Compared against primary clinician diagnosis of NPD, the MCMI-II-N scale had a sensitivity of 0.59 and a specificity of 0.96, which means that this scale was more successful in detecting who did not have NPD than detecting who did have this disorder.

Turley and colleagues (1992) also investigated the diagnostic concordance between the MCMI-II and the SIDP for 21 recent onset Bipolar disorder patients. Diagnostic frequency between the two measures was poor indicating a lack of agreement between the two measures concerning who has a personality disorder and who does not (see Table 4). The MCMI-II diagnosed 9 NPD cases, the SIDP diagnosed 2, and only 1 of these cases was diagnosed as NPD by both measures. Rate of personality disorder diagnoses was almost one-third higher for the MCMI-II when compared with the SIDP. Level of agreement between the MCMI-II and the SIDP on overall presence or absence of personality disorders was also poor. Comparisons for each personality disorder showed that the most dramatic diagnostic discrepancies between the two instruments were the high comparative frequencies of the narcissistic and antisocial personality disorder scales reported by the MCMI-II.

In a comparison of the concurrent validity of the MCMI-II and the PDE with a sample of 97 outpatients (Soldz et al., 1993), number of patients (66) assigned personality disorders by the MCMI-II was closer to the number of patients assigned PDE probable and definite diagnoses (75) than the number given to the definite diagnoses alone (43). The MCMI-II diagnosed 11 patients with a diagnosis of NPD whereas the PDE diagnosed only 2 patients as NPD. Overall diagnostic agreement between the MCMI-II and PDE on the presence or absence of any personality disorder was significantly related, as was the relationship between the two instruments in the assessment of cluster B personality disorders. However, agreement between these two instruments in the diagnosis of NPD was very low. Diagnostic efficiency statistics were equally as poor for the MCMI-II when based on PDE diagnoses for NPD. However, these classification rates improved in the diagnosis of cluster B personality disorders and for the presence or absence of any personality disorder. Agreement between the PDE and MCMI-II dimensional scores for NPD was also explored and the scores were found to be significantly related. However, the MCMI-II Schizoid, Histrionic, Passive-Aggressive, Paranoid, Antisocial, and Borderline scales were also significantly related to the dimensional PDE-NPD score, with the latter three scales having higher correlations than the NPD scale of the MCMI-II. It appears that the MCMI-II has high specificity and negative predictive power for all diagnoses, suggesting that the MCMI-II scales tend to agree on who does not have a personality disorder diagnosis. However, the low sensitivity and positive predictive power suggest that these two assessment measures will frequently disagree on who does have a specific personality disorder diagnosis.

Renneberg and colleagues (1992) examined the concurrent validity of personality disorder diagnosis of the SCID-II and MCMI-II. However, Kappa coefficients were not computed for NPD, because this disorder was not diagnosed 3 or more times in this sample. Instead, the Kappa coefficient for agreement of the MCMI-II (BR > 84) and

SCID-II personality disorder diagnoses of 54 agoraphobic outpatients for cluster B personality disorders was calculated and found to be significant.

The MCMI performed adequately when used as a screening instrument to identify the presence or absence of a personality disorder and for diagnosis by Axis II cluster. However, the MCMI may have only limited utility in identifying specific personality disorders. The MCMI personality disorder scales are generally poor in identifying patients who meet *DSM-III-R* criteria for corresponding disorders. Specifically, in assessing the comparison between the MCMI-II and structured interviews (PDE and SIDP), the MCMI tends to overdiagnose narcissistic personality disorder. This poor relationship in the diagnosis of NPD can also be seen in the relationship between MCMI-N scores and checklist ratings for NPD criteria. Despite such poor performance in categorical agreement, the dimensional relationship of the MCMI-N scale score was significantly correlated with SIDP-N. Also, the correlation between the MMPI-NPD scales and the NPI were found to be highly correlated with the MCMI-N. Finally, comparison of the narcissistic with nonnarcissistic patients on the MCMI revealed that Scale 5 (narcissism) was significantly greater than that of the nonnarcissistic group.

PROJECTIVE TECHNIQUES

Rorschach

Recent studies have used the Rorschach in differential diagnostic research for NPD, Table 5 summarizes results of these 6 studies. The first study, carried out by Farris (1988), found that compared to BPD patients, NPD patients showed significantly higher cognitive-perceptual functioning, responses indicative of body narcissism and phallic-oedipal issues. BPD patients produced significantly more splitting, projective identification (Lerner Defense Scales [LDS]; Lerner & Lerner, 1980), and primitive object representations. Berg (1990) similarly investigated the difference between NPD and BPD and found that the BPD group demonstrated significantly higher number of unusual percepts and splitting responses defined by the Rorschach Defense Scales (Cooper & Arnow, 1986), as well as less grandiosity compared with a NPD group.

Gacono et al. (1992) noted that psychopathic antisocial patients and NPDs produced a similar number of reflection responses, compared with nonpsychopathic antisocial patients, BPDs and Exner's nonpatient men (1990). The egocentricity ratio also yielded significant between-group and main effects. These authors also note the high numbers of personalized (PER) responses in both the psychopathic antisocial and narcissistic patients (3.32 and 2.00, respectively). In addition, borderline and psychopathic antisocial groups produced significantly more primitive object relations responses to the blots than the narcissistic patients, as well as a great deal of violent symbiotic separation and reunion themes in their content. Gacono et al. (1992) found that the narcissistic group produced a large number of idealization responses (56%), a finding similar to that of Hilsenroth et al. (1993). Narcissistic patients produced significantly more primitive idealization responses than both the psychopathic and nonpsychopathic antisocial groups. Also, NPD patients produced significantly more diffuse-shading (Y) and texture-shading (T) responses than the psychopathic antisocial comparison group.

Hilsenroth et al. (1993) further investigated differences in BPD and NPD with a clinical control group of cluster C personality disorders, by examining Rorschach content variables designed to assess defensive structures, aspects of aggression, and egocentricity. BPDs were found to employ primitive defensive structures (splitting and projective identification) to a greater degree and severity, were found to show more intense and

TABLE 5. Rorschach Variables Associated With the DSM Narcissistic Personality Disorder

Study	Variable ^a	NPD vs. Contrast Group ^b
Farris, 1988	Cognitive perceptual functioning	NPD > BPD
	Body narcissism	NPD > BPD
	Phallic symbols	NPD > BPD
	Primitive object representations	NPD < BPD
	Splitting	NPD < BPD
	Projective identification	NPD < BPD
Berg, 1990	Splitting	NPD < BPD
	Form quality unusual (FQu)	NPD < BPD
	Grandiosity	NPD > BPD
Gacono et al., 1992	Egocentricity	NPD > NP-ANPD
	Sum texture (T)	NPD > P-ANPD
	Sum diffuse shading (Y)	NPD > P-ANPD
	Idealization	NPD > NP-ANPD, P-ANPD
Hilsenroth et al., 1993	Primitive object representation	NPD < P-ANPD, BPD
	Splitting	NPD < BPD
	Projective identification	NPD < BPD
	Secondary process aggression	NPD < BPD
	Idealization	NPD > CC
Berg et al., 1993	Egocentricity	NPD > BPD
	Narcissistic object representations	NPD > Schiz.
	Primitive object representations	NPD < BPD
Hilsenroth et al., in press	Reflections	NPD > BPD, NP-ANPD, CC, CA, NC
	Egocentricity	NPD > NP-ANPD, NC
	Personalizations	NPD > NC
	Idealization	NPD > NP-ANPD, CA, CC, NC

Note. NPD = Narcissistic PD; BPD = Borderline PD; CA = DSM Cluster A PDs; CC = DSM Cluster C PDs; NC = Non Clinical group P-ANPD = Psychopathic Antisocial PDs; NP-ANPD = Non-psychopathic Antisocial PDs; Schiz. = Schizophrenic Group.

^aRorschach variables are defined in the body of the paper.

^bAll contrast comparisons significant at $p \leq .05$.

overall amounts of aggression. NPDs evinced significantly higher levels of egocentricity than borderlines and higher levels of idealization than the cluster C group.

Berg et al. (1993) found that borderline and narcissistic patients produced a significantly greater number of object relational scores representing figures in need of some external source of support than a schizophrenic group. For these responses the object is found to exist only insofar as it is an extension or reflection of another object. In comparison with NPDs, BPD patients produced significantly more object relational themes of severely imbalanced, malevolent, and engulfing relationships. Rorschach protocols of the narcissistic group reflected a difficulty in relating to others on a mutually autonomous basis, thereby relying solely on need satisfying relationships. These findings appear to highlight prior theoretical work concerning the character structure of those patients with disorders of the self as noted by Kohut (1971).

Recently, Hilsenroth et al. (in press) have investigated the extent to which the Rorschach is able to accurately identify pathological expressions of narcissism contrasting a sample of 91 patients (who were found to meet *DSM-IV* criteria for an Axis II disorder) and a control group of 50 nonclinical subjects on four Rorschach variables. These variables were: number of reflection (REF), personalized (PER), and idealization responses, as well as the egocentricity index (EGOI). Results of this study indicate that selected Rorschach variables can be used effectively to differentiate pathologically narcissistic patients from a nonclinical sample as well as from cluster A and cluster C personality disorders. The findings reported show that variables from the Rorschach can be used to aid in the differential diagnosis of NPD patients in relation to other personality disorders from within the *DSM-IV* B cluster diagnoses. Also, reflection and idealization variables were found to be empirically related to the MMPI-2-NPD-NO scale (both at $r = 0.30$, $p < 0.02$). In addition, number of reflection responses that a patient produced on his/her Rorschach protocol was significantly and positively related ($r = 0.33$, $p < 0.003$) to the patient's total number of *DSM-IV* criteria for NPD. With regard to individual *DSM-IV* NPD criteria, production of a reflection response was found to be associated with fantasies of unlimited success, sense of entitlement, and a grandiose sense of self importance. It is interesting to note that this variable, derived from a projective test, was significantly related to *DSM-IV* NPD criteria associated with the intrapsychic or cognitive features pertaining to pathological narcissism moreso than to behavioral expressions. Although the idealization response score was not significantly related to total number of *DSM-IV* NPD criteria, it was related to NPD criterion 2 (fantasies of unlimited success). The relationship of these two scores to the more intrapsychic or internal characteristics of NPD suggest that the Rorschach may prove to be very useful when employed in tandem with other methods of assessment that are designed to assess more overt/behavioral expressions of NPD. Finally, these variables (reflection and idealization) from the Rorschach could also be employed for classification purposes in ways that were clinically meaningful in the diagnosis of NPD.

It appears from past research that the Rorschach may be helpful to identify NPD patients from different clinical groups. In relation to BPDs narcissistic patients develop higher levels of object representations, employ less primitive and severe defenses (i.e., splitting and projective identification) and less aggressive imagery. NPDs will also tend to develop more reflection responses than BPD patients or a majority of nonpsychopathic ANPDs and more primitive idealized responses than this same group of ANPDs and cluster C personality disorders.

THEMATIC APPERCEPTION TEST (TAT)

A few studies have sought to assess the ability of the Thematic Apperception Test (TAT) in the assessment of a narcissistic character style. The first of these was the work of Leary (1957), who utilized TAT stories to aid in the determination of an interpersonal style. Leary (1957) categorized the narratives from these stories using 16 codes, 2 of which he labeled as *narcissistic* and *exploitive*. Stories met these criteria if the interaction could be labeled as one of independence, struggle over power, selfishness, seduction, etc. Whereas Leary (1957) presented data for different clinical groups, his work pre-dated the current diagnostic taxonomic systems represented in the *DSM-III/DSM-III-R/DSM-IV*. Similarly, Harder (1979) presented a scale designed to assess an ambitious-narcissistic character style on the TAT. In regard to scoring this scale, the interrater agreement and reliability coefficients were high. Also, TAT scores

were found to be significantly related to similar scales on a set of early memories and the Rorschach. In addition, TAT scale ratings were also found to significantly differentiate subjects rated as ambitious from those rated as nonambitious. Whereas this scale has shown reliability and validity in assessing an ambitious–narcissistic character style, the study utilized “40 relatively well-functioning” male college students, which may limit the generalizability of these findings to a pathological group comprised of NPDs. This same criticism, the exclusive use of undergraduates for subjects, may also be applied to the Narcissism-Projective which asks subjects to describe two TAT cards and two early childhood memories (Shulman & McCarthy, 1986). Responses to these stimuli are scored employing criteria based on the *DSM-III* for NPD. Scores on this measure have been found to have high interrater reliability coefficients and to be significantly related to interview ratings of behavior, as well as an objective measure of narcissism, the NPI (Shulman, McCarthy, Ferguson, 1988; Shulman & Ferguson, 1988). While both of these TAT scales have shown utility in the assessment of a narcissistic construct, further research must be conducted to explore what contribution they may make to differential diagnosis and diagnostic efficiency. To date, the TAT has been primarily used at an ideographic level to understand individual NPD patients. Nomothetic data utilizing this test in differential diagnosis may prove useful as well. Further research might begin to attempt to provide some large sample norms in the differential diagnosis of Axis II disorders.

Whereas several psychiatric conditions have been extensively studied using the TAT, few of these studies have examined the test characteristics of character pathology (Bellak, 1986). Over the last decade, there has been some investigation of thematic test analysis of narcissistic personality disorders but most have been case studies (Abrams, 1993) and many of these appear in the international literature (Brelet, 1981, 1983, 1986, 1987; Shentoub et al., 1990; Seifert, 1984). Recently, Brelet (1994) presented a multidimensional interpretative model for narcissistic narratives in the TAT for clinical patients, but this work has yet to receive any empirical support.

INTEGRATION

Our review is intended to be helpful to clinicians and researchers alike who need to evaluate the different methods available for assessing pathological narcissism. In this section, we integrate and summarize the material previously reviewed. Semi-structured interviews have proven to be a fairly reliable and valid method for identifying Axis II disorders in general, but for the most part, the studies reviewed contained inadequate samples of NPD subjects, making it more difficult to evaluate their ability to identify this particular personality disorder. The one exception to this general limitation is Gunderson and Ronningstam's (1987) DIN, which was specifically developed to assess a wide range (DSM and beyond) of traits making up narcissistic pathology. Therefore, the researcher or clinician interested primarily in identifying a wide range of criteria or traits associated with NPD, may find this specialized assessment instrument optimal.

Self-report measures for the DSM personality disorders appear to be best at screening for the presence of absence of any personality disorder and are less useful for identifying specific personality disorders such as NPD. In fact, the generally high specificity and negative predictive power (see Table 4) of these instruments indicates that their strength is in identifying patients who are true negatives for NPD. From this review, it appears that self-report tests are most useful as quick and economical screening tests for NPD or other DSM personality disorders.

Despite satisfactory reliability for both self-report inventories and structured interviews, the present review revealed only limited support for the concurrent validity of these methods to diagnosis NPD. Historically, concurrent validity has been difficult to demonstrate for any method of diagnosing the DSM personality disorders (Hurt et al., 1984; Edell, 1984; Zimmerman, 1994). Other recent reviews have also documented poor to moderate agreement across personality disorder diagnostic methods (Perry, 1992; Zimmerman, 1994). In part, this is due to the fact that self-report measures diagnose personality disorders at a much higher frequency than do clinicians or structured interviews. Moreover, self-report measures frequently result in multiple personality disorder diagnoses being attributed to an individual. This review shows that these general concerns regarding self-report measures are also quite relevant to the NPD diagnosis. Self-report inventories are far more likely to diagnose NPD and to indicate the presence of multiple personality disorders than are structured interviews.

In addition, self-report measures have shown poor categorical agreement across the 11 personality disorders, including NPD. Whereas agreement on categorical diagnoses has been poor, these instruments have shown a moderate degree of correlation when the narcissistic criteria are dimensionalized (see Table 2 and Table 3). Structured interviews identified NPD as much less prevalent than did the self-report inventories. Disagreement between these two diagnostic methods may be due to the high levels of overlap of the HPD, ANPD, and BPD with the NPD scales on the self-report measures. This finding may also reflect an exaggeration of symptom severity in the use of self-report measures. It appears that utilizing self-report inventories as the sole criterion for making *DSM-III-R/IV* diagnoses of NPD is inadvisable.

We must conclude that at all levels, self-report inventories, and the structured interviews are often in disagreement concerning the presence of specific personality pathology. Therefore, these measures cannot be used interchangeably and attention should be paid in literature reviews of personality disorders to the examination of the effects that may be due to the diagnostic method that is employed. However, note that employing the more conservative base rate > 84 as is done with the MCMI-II, can improve the agreement rates between structured interviews and self-report measures (Chick et al., 1993; Renneberg et al., 1992; Torgerson & Alnaes, 1990).

From the research just reviewed, it is clear that self-report inventories are more likely to give an overabundance of false positive diagnoses compared with ratings by clinicians or a structured interview. Thus, self-report instruments provide a broad band of personality disorder data and are well-suited to preliminary exploration of the clinical picture. To their credit, Hyler, Skodol, Kellman, Oldham, and Rosnick (1990) suggested that the PDQ-R should be used to screen individuals to be studied further for specific personality disorders; they recommended that positive diagnoses according to the PDQ-R be verified for clinical significance by clinician-administered interviews. We endorse this approach and believe that it should be used with all self-report measures of NPD.

The work of Chatham and colleagues (1993) has been one of the two attempts to specifically assess the construct validity of the MMPI-NPD scale and the MCMI-N in the assessment of pathological narcissism. It appeared that these two measures have utility in the differentiation of those patients with narcissistic pathology from those without such pathology. However, ability of these scales to evaluate questions about differential diagnosis of NPD from related subgroups remains unanswered. Whereas both measures performed very well, the narcissistic group's high scores on scales 4 and 6 of the MCMI again raise the specter of poor differential diagnosis and multiple diagnosis of related personality disorders when the MCMI is used in isolation. In addition, it would have been useful if the authors had presented values for the MMPI BPD,

HPD, and ANPD scales on both groups (narcissistic vs. nonnarcissistic patients) to evaluate the differential validity of these MMPI-PD scales. The second study by Castlebury and colleagues (in press) found good convergent validity between the MMPI-2-NPD scales and *DSM-IV* criteria, but the diagnostic efficiency statistics calculated for the classification of the NPD patients were unremarkable.

It appears that interviews allow for greater flexibility in the assessment of NPD because clinical judgment may be necessary to determine or clarify whether the diagnostic aspects of a patient's behavior are present (e.g., *DSM-IV* Criterion 9: Arrogant and haughty behaviors). In addition, observation of pathological behavior may be more useful to a clinician than the sole reliance on self-report. Past authors have criticized self-report inventories concerning the assessment of NPD because these instruments tend to be more direct in identifying narcissistic traits and therefore are more likely to evoke defensive responses (Gunderson et al., 1990). Moreover, these authors also state that NPD patients are particularly unable to view themselves in a realistic manner. Although allowing for the clinical observation of behavior, one has to wonder if this same criticism might also apply, at least in part, to semi-structured interviews. Additionally, interviews have limitations that clinicians should be aware of as well. Past research has also indicated that clinicians underestimate or minimize co-existing syndromes once the presence of one or two Axis II disorders have been recognized (Widiger & Frances, 1987). Furthermore, unlike self-report inventories, which may include indices to detect intentional response dissimulation (faking), exaggeration of symptoms, random responding, acquiescence or denial, clinical interviewers may be susceptible to active attempts at malingering. As such, assessment of personality disorder criteria may be difficult through direct inquiry and it is questionable whether NPD patients would admit that they are egocentric, self-indulgent, inconsiderate, or interpersonally exploitive. Also, with the advent of the *DSM-IV* it will be necessary for at least some limited revision to be undertaken for a number of these interview and self-report inventories to remain consistent with standardized psychiatric classification.

Given the concerns that have been presented for limitations of NPD patients reporting diagnostic criteria, either on self-report measures or semi-structured interviews, it is surprising that projective techniques have been under-utilized. Furthermore, research over the last decade has shown this mode of assessment to be useful in the differential diagnosis of NPD from both related and unrelated personality disorders. However, only recently has research been conducted that has reported information concerning reliability of the NPD diagnosis, assesses the relationship between *DSM* criteria and projective test variables, or calculates the diagnostic efficiency statistics for projective variables with regard to correct classification (Hilsenroth et al., in press). Future research with projective test variables should continue to address these issues by employing more stringent methodological standards. It will be important to ascertain the contribution of projective techniques to the diagnosis of NPD in comparison to and in conjunction with what the semi-structured interview and self-report inventory methods have to offer.

CONCLUSIONS

The importance of a multi-method approach to assessment has been stressed by several different authors (Rappaport, Gill & Schafer, 1968; Leary, 1957; Campbell & Fiske, 1959; Jackson, 1971). Implicit in this approach is the idea that individuals are multi-dimensional beings who vary not only from one another but also in the way oth-

ers view them (social perception), the way they view themselves (self perception), and the ways in which underlying dynamics will influence their behavior (intentionality/fantasy/ideals). Such an approach presents clinicians and researchers with the responsibility to sample each domain of functioning. This form of assessment may aid clinicians in obtaining a comprehensive understanding of an individual rather than focusing on just one facet of behavior. Also, the concept of narcissism has frequently been viewed as multidimensional and, therefore, a single score on any one measure may be far less optimal than an assessment process which provides information concerning the multiple aspects of narcissistic pathology (i.e., grandiosity, need for mirroring/admiration, narcissistic rage, entitlement, etc).

In conclusion, it seems prudent to encourage clinicians and researchers alike to employ multiple methods of assessing narcissism and to utilize this information in a systematic and theoretically consistent fashion. Understanding the variety of options available for the measurement of narcissistic pathology is useful in the comparison and selection of the various methods and techniques. Future research should attempt to compare the diagnostic efficiency of those methods just reviewed to determine their ability, alone and in combination to accurately classify NPD patients. It may be that these approaches in combination are more effective than each considered separately, or conversely, that there may be relative advantages of each method that are unique to an assessment of NPD. Due to the recent introduction of the *DSM-IV*, a continued evaluation of the convergent and discriminant validity of these three approaches to assessment of NPD is needed.

Acknowledgement — Some portion of this article was completed by Mark J. Hilsenroth while a Clinical Fellow at The Cambridge Hospital and Harvard Medical School.

REFERENCES

- Abrams, D. (1993). Pathological narcissism in an eight-year-old boy: An example of Bellak's TAT and CAT diagnostic system. *Psychoanalytic Psychology, 10*, 573-591.
- American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders (3rd ed.)*. Washington, DC: Author.
- American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders (3rd ed., rev.)*. Washington, DC: Author.
- Ames-Frankel, J., Devlin, M., Walsh, B., Strasser, T., Sadik, C., Oldham, J., & Roose, S. (1992). Personality disorder diagnosis in patients with bulimia nervosa: Clinical correlates and changes with treatment. *Journal of Clinical Psychiatry, 53*, 90-96.
- Bellak, L. (1986). *The TAT, CAT, and SAT in clinical use (4th ed., rev.)*. Orlando, FL: Grune & Stratton.
- Berg, J. (1990). Differentiating ego functions of borderline and narcissistic personalities. *Journal of Personality Assessment, 55*, 537-548.
- Berg, J., Packer, A., & Nunno, V. (1993). A Rorschach analysis: Parallel disturbance in thought and in self/object representation. *Journal of Personality Assessment, 61*, 311-323.
- Berg, M. (1983). Borderline psychopathology as displayed on psychological tests. *Journal of Personality Assessment, 47*, 120-123.
- Blashfield, R., Blum, N., & Pfohl, B. (1992). The effects of changing axis II diagnostic criteria. *Comprehensive Psychiatry, 33*, 245-252.
- Blatt, S., & Lerner, H. (1983). The psychological assessment of object representations. *Journal of Personality Assessment, 47*, 7-28.
- Brelet, F. (1981). On narcissism in the TAT. *Psychologie Francaise, 26*, 24-37.
- Brelet, F. (1983). The TAT and narcissism: Dynamic and economic perspectives. *Psychologie Francaise, 28*, 119-123.
- Brelet, F. (1986). *The TAT: Fantasy and the projective situation. Narcissism, borderline functioning, depression*. Paris: Dunod.

- Brelet, F. (1987). One seeks a director. *Psychologie Francaise*, 32, 137-140.
- Brelet, F. (1994). Expression of narcissistic fantasy in the TAT. *Rorschachiana*, 19, 97-111.
- Brent, D., Zelenak, J., Bukstein, O., & Brown, R. (1990). Reliability and validity of the Structured Interview for personality disorders in adolescents. *Journal of the American Academy for Child and Adolescent Psychiatry*, 29, 349-354.
- Brooks, R., Baltazar, P., McDowell, D., Munjack, D., & Bruns, J. (1991). Personality disorders co-occurring with panic disorder with agoraphobia. *Journal of Personality Disorders*, 5, 328-336.
- Campbell, D., & Fiske, D. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81-105.
- Castlebury, F., Hilsenroth, M., Handler, L., & Durham, T. (in press). Use of the MMPI-2 personality disorder scales in the assessment of DSM-IV antisocial, borderline and narcissistic personality disorders. *Assessment*.
- Chatham, P., Tibbals, C., & Harrington, M. (1993). The MMPI and the MCMI in the evaluation of narcissism in a clinical sample. *Journal of Personality Assessment*, 60, 239-251.
- Chessick, R. (1987). *Narcissism and the sense of self*. New York: Aronson.
- Chick, D., Sheaffer, C., Goggin, W., & Sison, G. (1993). The relationship between MCMI personality scales and clinician generated DSM-III-R personality disorder diagnoses. *Journal of Personality Assessment*, 61, 264-276.
- Colligan, R., Morey, L., & Offord, K. (1994). The MMPI/MMPI-2 personality disorder scales: Contemporary norms for adults and adolescents. *Journal of Clinical Psychology*, 50, 168-200.
- Davis, R., Blashfield, R., & McElroy, R. (1993). Weighting criterion in the diagnosis of a personality disorder: A demonstration. *Journal of Abnormal Psychology*, 102, 319-322.
- Dowson, J. (1992). DSM-III-R narcissistic personality disorder evaluated by patients' and informants' self-report questionnaires: Relationship with other personality disorders and a sense of entitlement as an indicator of narcissism. *Comprehensive Psychiatry*, 33, 397-406.
- Dubro, A., & Wetzler, S. (1989). An external validity study of the MMPI personality disorder scales. *Journal of Clinical Psychology*, 45, 570-575.
- Dubro, A., Wetzler, S., & Kahn, M. (1988). A comparison of three self-report questionnaires for the diagnosis of DSM-III personality disorders. *Journal of Personality Disorders*, 2, 256-266.
- Edell, W. (1984). The borderline syndrome index: Clinical validity and utility. *Journal of Nervous and Mental Disease*, 172, 254-263.
- Farris, M. (1988). Differential diagnosis of borderline and narcissistic personality disorders. In H. Lerner & P. Lerner (Eds.), *Primitive mental states and the Rorschach* (pp. 299-338). New York: International Universities Press.
- First, M., Spitzer, R., Gibbon, M., Williams, J., Davies, M., Borus, J., Howes, M., Kane, J., Pope, H., & Rounsaville, B. (1995). The Structured clinical interview for DSM-III-R personality disorders (SCID-II), II: Multisite test-retest reliability study. *Journal of Personality Disorders*, 9, 92-104.
- Gacono, C., Meloy, J., & Heaven, T. (1990). A Rorschach investigation of attachment and anxiety in antisocial personality disorder. *Journal of Personality Assessment*, 55, 270-279.
- Gacono, C., Meloy, J., & Berg, J. (1992). Object relations, defensive operations, and affective states in narcissistic, borderline and antisocial personality disorder. *Journal of Personality Assessment*, 59, 32-49.
- Gunderson, J., & Ronningstam, E. (1987). *The Diagnostic interview for narcissistic patients*. Belmont, MA: McLean Hospital.
- Gunderson, J., Ronningstam, E., & Bodkin, A. (1990). The diagnostic interview for narcissistic patients. *Archives of General Psychiatry*, 47, 676-680.
- Gunderson, J., Ronningstam, E., & Smith, L. (1991). Narcissistic personality disorder: A review of data on DSM-III-R descriptions. *Journal of Personality Disorders*, 5, 167-177.
- Guthrie, P. C., & Mobley, B. D. (1994). A comparison of the differential diagnostic efficiency of three personality disorder inventories. *Journal of Clinical Psychology*, 50, 656-664.
- Harder, D. (1979). The assessment of ambitious-narcissistic character style with three projective tests: The early memories, TAT, and Rorschach. *Journal of Personality Assessment*, 43, 23-32.
- Hills, H. (1995). Diagnosing personality disorders: An examination of the MMPI-2 and MMCI-II. *Journal of Personality Disorders*, 65, 21-34.
- Hilsenroth, M., Fowler, C., Padawer, J., & Handler, L. (in press). Narcissism in the Rorschach revisited: Some reflections upon empirical data. *Psychological Assessment*.
- Hilsenroth, M., Hibbard, S., Nash, M., & Handler, L. (1993). A Rorschach study of narcissism, defense, and aggression in borderline, narcissistic and cluster C personality disorders. *Journal of Personality Assessment*, 60, 346-361.
- Hogg, B., Jackson, H., Rudd, R., & Edwards, J. (1990). Diagnosing personality disorders in recent-onset schizophrenia. *Journal of Nervous and Mental Disease*, 178, 194-199.
- Hunt, C., & Andrews, G. (1992). Measuring personality disorder: The use of self report questionnaires. *Journal of Personality Disorders*, 6, 125-133.

- Hurt, S., Clarkin, J., & Morey, L. (1990). An examination of the stability of the MMPI personality disorder scales. *Journal of Personality Assessment*, 54, 16–23.
- Hyler, S., & Lyons, M. (1988). Factor analysis of the DSM–III personality clusters: A replication. *Comprehensive Psychiatry*, 29, 304–308.
- Hyler, S., & Reider, R. (1987). *Personality Disorder Questionnaire—Revised: PDQ-R*. New York: New York State Psychiatric Institute.
- Hyler, S., Rieder, R., Williams, J., Spitzer, R., Lyons, M., & Hendler, J. (1988). The Personality Diagnostic Questionnaire: Development and preliminary results. *Journal of Personality Disorders*, 2, 229–237.
- Hyler, S., Reider, R., Spitzer, R., & Williams, J. (1982). *The Personality Diagnostic Questionnaire (PDQ)*. New York: New York State Psychiatric Institute.
- Hyler, S., Rieder, R., Williams, J., Spitzer, R., Lyons, M., & Hendler, J. (1989). A comparison of clinical and self-report diagnoses of DSM–III personality disorders in 552 patients. *Comprehensive Psychiatry*, 30, 170–178.
- Hyler, S., Skodol, A., Kellman, D., Oldham, J., & Rosnick, L. (1990). Validity of the Personality Disorder Questionnaire-Revised: Comparison with two structured interviews. *American Journal of Psychiatry*, 147, 1043–1048.
- Jackson, D. (1971). The dynamics of structured personality tests. *Psychological Review*, 78, 229–248.
- Kernberg, O. (1970). Factors in the psychoanalytic treatment of narcissistic personalities. *Journal of the American Psychoanalytic Association*, 18, 51–85.
- Kernberg, O. (1975). *Borderline conditions and pathological narcissism*. New York: Aronson.
- Kernberg, O. (1984). *The treatment of severe character disorders*. New Haven, CT: Yale University Press.
- Kohut, H. (1971). *The analysis of the self*. New York: International Universities Press.
- Kohut, H. (1977). *The restoration of the self*. New York: International Universities Press.
- Leary, T. (1957). *Interpersonal diagnosis of personality*. New York: Ronald.
- Lerner, H. (1988). The narcissistic personality as expressed through psychological tests. In H. Lerner & P. Lerner (Eds.), *Primitive mental states and the Rorschach* (pp. 257–298). New York: International Universities Press.
- Lerner, P., & Lerner, H. (1980). Rorschach assessment of primitive defenses in borderline personality structure. In J. Klawer, H. Lerner, P. Lerner, & A. Sugarman (Eds.), *Borderline phenomena and the Rorschach test* (pp. 257–274). New York: International Universities Press.
- Loranger, A. (1988). *Personality Disorder Examination (PDE) manual*. Yonkers, NY: DV Communications.
- Loranger, A., Lenzweger, M., Gartner, A., Susman, V., Herzig, J., Zammit, G., Gartner, J., Abrams, R., & Young, R. (1991). Trait-state artifacts and the diagnosis of personality disorders. *Archives of General Psychiatry*, 48, 720–728.
- Loranger, A., Oldham, J., & Tullis, E. (1982). Familial transmission of DSM-III borderline personality disorder. *Archives of General Psychiatry*, 39, 795–799.
- Loranger, A., Susman, V., Oldham, J., & Russakoff, L. (1987). The Personality Disorder Examination: A preliminary report. *Journal of Personality Disorders*, 1, 1–13.
- Masterson, J. (1981). *The narcissistic and borderline disorders*. New York: Brunner/Mazel.
- Miller, H., Streiner, D., & Parkinson, A. (1992). Maximum likelihood estimates of the ability of the MMPI and MCMI personality disorder scales and the SIDP to identify personality disorders. *Journal of Personality Assessment*, 6, 1–13.
- McCann, J. (1990). Bias and MCMI-II diagnosis. *Journal of Psychopathology and Behavioral Assessment*, 12, 17–26.
- McCann, J. (1989). MMPI personality disorder scales and the MCMI: Concurrent validity. *Journal of Clinical Psychology*, 45, 365–369.
- McCann, J. (1991). Convergent and discriminant validity of the MCMI-II and the MMPI personality disorder scales. *Psychological Assessment*, 3, 9–18.
- McGlashan, T., & Heinssen, R. (1989). Narcissistic, antisocial, and noncomorbid subgroups of borderline disorder. *Psychiatric Clinics of North America*, 12, 653–670.
- Meissner, W. (1978). Narcissistic personalities and borderline conditions: A differential diagnosis. *Annual of Psychoanalysis*, 7, 171–201.
- Miller, H. R., Streiner, D. L., & Parkinson, A. (1992). Maximum likelihood estimates of the ability of the MMPI and MCMI personality disorder scales and the SIDP to identify personality disorders. *Journal of Personality Assessment*, 6, 1–13.
- Millon, T. (1981). *Disorders of personality: DSM–III, Axis II*. New York: Wiley.
- Millon, T. (1983). *Millon Clinical Multiaxial Inventory manual*. Minneapolis: National Computer Systems.
- Millon, T. (1987). *Millon Clinical Multiaxial Inventory manual* (2nd ed.). Minneapolis, MN: National Computer Systems.
- Morey, L. (1985). A psychometric analysis of five DSM–III categories. *Personality and Individual Differences*, 6, 323–329.

- Morey, L. (1988). The categorical representation of personality disorder: A cluster analysis of DSM-III-R personality features. *Journal of Abnormal Psychology, 97*, 314-321.
- Morey, L., & Le Vine, D. (1988). A multitrait-multimethod examination of the MMPI and MCMI. *Journal of Psychopathology and Behavioral Assessment, 10*, 333-344.
- Morey, L., Waugh, M., & Blashfield, R. (1985). MMPI scales for DSM-III personality disorders: Their derivation and correlates. *Journal of Personality Assessment, 49*, 245-251.
- Oldham, J., Skodol, A., Kellman, D., Hyler, S., Rosnick, L., & Davies, M. (1992). Diagnosis of DSM-III-R personality disorders by two structured interviews: Patterns of comorbidity. *American Journal of Psychiatry, 149*, 213-220.
- Perry, J. (1992). Problems and considerations in the valid assessment of personality disorders. *American Journal of Psychiatry, 149*, 1645-1652.
- Pilkonis, P., Heape, C., Ruddy, J., & Serrao, P. (1991). Validity in the diagnosis of personality disorders: The use of the LEAD standard. *Psychological Assessment, 3*, 46-54.
- Pfohl, B., Blum, N., Zimmerman, M., & Stangl, D. (1989). *Structured interview for DSM-III personality disorders (rev. ed.) (SIDP-R)*. University of Iowa, College of Medicine, Iowa City, Iowa 52240.
- Pfohl, B., Coryell, W., Zimmerman, M., & Stangl, D. (1986). DSM-III personality disorders: Diagnostic overlap and internal consistency of individual DSM-III criteria. *Comprehensive Psychiatry, 27*, 21-34.
- Pfohl, B., Stangl, D., & Zimmerman, M. (1983). *Structured interview for DSM-III personality disorders (SIDP)* (2nd ed.). University of Iowa, College of Medicine, Iowa City, Iowa 52240.
- Piersma, H. (1986). The stability of the Millon Clinical Multiaxial Inventory for psychiatric inpatients. *Journal of Personality Assessment, 50*, 193-197.
- Plakun, E. (1987). Distinguishing narcissistic and borderline personality disorders using the DSM-III criteria. *Comprehensive Psychiatry, 28*, 437-443.
- Plakun, E. (1989). Narcissistic personality disorder: A validity study and comparison to borderline personality disorder. *Psychiatric Clinics of North America, 12*, 603-620.
- Pope, H., Jonas, J., Hudson, J., Cohen, B., & Gunderson, J. (1983). The validity of DSM-III borderline personality disorder. *Archives of General Psychiatry, 40*, 23-30.
- Prifitera, A., & Ryan, J. (1984). Validity of the narcissistic personality inventory in a psychiatric sample. *Journal of Clinical Psychology, 40*, 140-142.
- Pulver, S. (1970). Narcissism: The term and the concept. *Journal of the American Psychoanalytic Association, 18*, 319-341.
- Rapaport, D., Gill, M., & Schafer, R. (1968). *Diagnostic psychological testing (rev. ed; R. Holt Ed.)*. New York: International Universities Press.
- Raskin, R., & Hall, C. (1979). A narcissistic personality inventory. *Psychological Reports, 45*, 596.
- Reich, J., & Troughton, E. (1988). Comparison of frequency of DSM-III personality disorders in panic outpatients and normal populations. *Psychiatric Research, 26*, 89-100.
- Renneberg, B., Chambless, D., Dowdall, D., Fauerbach, J., & Gracely, E. (1992). The Structured Clinical Interview for DSM-III-R, Axis I and the Millon Clinical Multiaxial Inventory: A concurrent validity study of personality disorders among anxious outpatients. *Journal of Personality Disorders, 6*, 117-124.
- Ronningstam, E. (1988). Comparing three systems for diagnosing narcissistic personality disorder. *Psychiatry, 51*, 300-311.
- Ronningstam, E., & Gunderson, J. (1988). Narcissistic traits in psychiatric patients. *Comprehensive Psychiatry, 29*, 545-549.
- Ronningstam, E., & Gunderson, J. (1990). Identifying criteria for narcissistic personality disorder. *American Journal of Psychiatry, 147*, 918-922.
- Ronningstam, E., & Gunderson, J. (1991). Differentiating borderline personality disorder from narcissistic personality disorder. *Journal of Personality Disorders, 5*, 225-232.
- Schuler, C., Snibbe, J., & Buckwalter, J. (1994). Validity of the MMPI personality disorder Scales. *Journal of Clinical Psychology, 50*, 220-227.
- Shentoub, V., Azoulay, C., Bailly-Salin, M., Benfredj, K., Boekholt, M., Brelet, F., Chabert, C., Chretien, M., Emmanuelli, M., Martin, M., Monin, E., Peruchon, M., & Serviere, A. (1990). *Manual for the use of the TAT: Psychoanalytic approach*. Paris: Dunod.
- Shulman, D., & Ferguson, G. (1988). Two methods of assessing narcissism: Comparison of the Narcissism-Projective (N-P) and the Narcissistic Personality Inventory (NPI). *Journal of Clinical Psychology, 44*, 857-866.
- Shulman, D., & McCarthy, E. (1986). *The projective assessment of narcissism: Manual for the N-P*. Unpublished Manuscript.
- Shulman, D., McCarthy, E., & Ferguson, G. (1988). The projective assessment of narcissism: Development, reliability, and of the N-P. *Psychoanalytic Psychology, 5*, 285-297.
- Seifert, W. (1984). *Character and its psychodiagnostic history with the TAT*. Munich/Basel: Ernst Reinhardt Verlag.

- Siever, L., & Klar, H. (1986). A review of DSM-III criteria for the personality disorders. In A. Frances & R. Hales (Eds.), *American Psychiatric Association annual review* (Vol. 5, pp. 299–301). Washington, DC: American Psychiatric Press.
- Skodol, A., Rosnick, L., Kellman, D., Oldham, J., & Hyler, S. (1988). Validating structured DSM-III-R personality disorder assessments with longitudinal data. *American Journal of Psychiatry*, *145*, 1297–1299.
- Soldz, S., Budman, S., Demby, A., & Merry, J. (1993). Diagnostic agreement between the Personality Disorder Examination and the MCMI-II. *Journal of Personality Assessment*, *60*, 486–499.
- Spitzer, R. (1983). Psychiatric diagnosis: Are clinicians still necessary? *Comprehensive Psychiatry*, *24*, 399–411.
- Spitzer, R., Williams, J., & Gibbon, M. (1987). *Instruction manual for the Structured Clinical Interview for DSM-III-R* (SCID). New York: Biometrics Research Department, New York State Psychiatric Institute.
- Standage, K., & Ladha, N. (1988). An examination of the reliability of the Personality Disorders Examination and a comparison with other methods of identifying personality disorders in a clinical sample. *Journal of Personality Disorders*, *2*, 267–271.
- Stangl, D., Pfohl, B., Zimmerman, M., Bowers, W., & Corenthal, C. (1985). A structured interview for the DSM-III personality disorders: A preliminary report. *Archives of General Psychiatry*, *42*, 591–596.
- Stanley, M., Turner, S., & Borden, J. (1990). Schizotypal features in obsessive-compulsive disorder. *Comprehensive Psychiatry*, *31*, 511–518.
- Stone, M. (1989). Long-term follow-up of narcissistic/borderline patients. *Psychiatric Clinics of North America*, *12*, 621–641.
- Stolorow, R. (1975). Toward a functional definition of narcissism. *International Journal of Psycho-Analysis*, *56*, 179–185.
- Streiner, D., & Miller, H. (1988). Validity of MMPI scales for DSM-III personality disorders: What are they measuring? *Journal of Personality Disorders*, *2*, 238–242.
- Teicholz, J. (1978). A selective review of the psychoanalytic literature on theoretical conceptualizations of narcissism. *Journal of American Psychoanalytic Association*, *26*, 831–861.
- Torgersen, S., & Alnaes, R. (1990). The relationship between MCMI personality scales and the DSM-III, Axis II. *Journal of Personality Assessment*, *55*, 698–707.
- Trull, T. (1992). DSM-III-R personality disorders and the five-factor model of personality: An empirical comparison. *Journal of Abnormal Psychology*, *101*, 553–560.
- Trull, T. (1993). Temporal stability and validity of two personality disorder inventories. *Psychological Assessment*, *5*, 11–18.
- Trull, T., & Goodwin, A. (1993). Relationship between mood changes and the report of personality disorder symptoms. *Journal of Personality Assessment*, *61*, 99–111.
- Turley, B., Bates, G., Edwards, J., & Jackson, H. (1992). MCMI-II personality disorders in recent onset bipolar disorders. *Journal of Clinical Psychology*, *48*, 320–329.
- Vaillant, G., & Drake, R. (1985). Maturity of ego defenses in relation to DSM-III axis II personality disorder. *Archives of General Psychiatry*, *42*, 597–601.
- Vaillant, G., & Perry, J. (1985). Personality disorders. In H. Kaplan & B. Sadock (Eds.), *Comprehensive textbook of psychiatry*, vol. 1 (4th ed). Baltimore, MD: Williams & Wilkins.
- Wechsler, D. (1981). *Wechsler Adult Intelligence Scale-Revised*. New York: The Psychological Corporation.
- Wetzler, S., & Dubro, A. (1990). Diagnosis of personality disorders by the Millon Clinical Multiaxial Inventory. *Journal of Nervous and Mental Disease*, *178*, 261–263.
- Westen, D. (1990). Towards a revised theory of borderline object relations: Contributions of empirical research. *International Journal of Psycho-Analysis*, *71*, 661–693.
- Widiger, T., & Frances, A. (1987). Interviews and inventories for the measurement of personality disorders. *Clinical Psychology Review*, *7*, 49–75.
- Widiger, T., Freiman, K., & Bailey, B. (1990). Convergent and discriminant validity of personality disorder prototypic acts. *Psychological Assessment*, *2*, 107–113.
- Widiger, T., Trull, T., Hurt, S., Clarkin, J., & Frances, A. (1987). A multidimensional scaling of the DSM-III personality disorders. *Archives of General Psychiatry*, *44*, 557–563.
- Wink, P., & Gough, H. (1990). New narcissism scales for the California Psychological Inventory and the MMPI. *Journal of Personality Assessment*, *54*, 446–462.
- Wise, E. A. (1996). Comparative validity of MMPI-2 and MCMI-II personality disorder classifications. *Journal of Personality Assessment*, *66*, 569–582.
- Yeung, A., Lyons, M., Waternaux, C., Faraone, S., & Tsuang, M. (1993). Empirical determination of thresholds for case identification: Validation of the PDQ-R. *Comprehensive Psychiatry*, *34*, 384–391.
- Zimmerman, M. (1994). Diagnosing personality disorders: A review of issues and research methods. *Archives of General Psychiatry*, *51*, 225–245.
- Zimmerman, M., & Coryell, W. (1990). Diagnosing personality disorders in the community: A comparison of self-report and interview measures. *Archives of General Psychiatry*, *47*, 527–531.