Evaluating Psychological Insight in a Clinical Sample Using the Shedler-Westen Assessment Procedure

Meaghan E. Lehmann, MA and Mark J. Hilsenroth, PhD

Abstract: Using the Shedler-Westen Assessment Procedure (SWAP)-200 Q-Sort items in a large clinical sample of outpatients (*N* = 105), we developed the SWAP insight scale. The rationale, psychometric properties, and convergent validity of this insight scale are reported. Through factor analysis, six of the SWAP-200 items were identified as psychometrically optimal in the assessment of insight (presence or absence) with an alpha coefficient of 0.78. We examined the construct validity of this SWAP insight scale using independent clinical videotape ratings of the Capacity for Dynamic Process Scale, specifically the items "appears introspective" and "manifests insight." We also examined the relationship between the SWAP insight scale using independent clinical videotape ratings of the Social Cognition and Object Relations Scale–Global Ratings, specifically ratings of "complexity of representations" and "social causality." The results demonstrated significant positive correlation between the SWAP insight scale and all five of these criterion measures. Partial correlations demonstrated that, even when the effects of global psychiatric severity are controlled for, the SWAP insight scale maintains a significant relationship with independent clinical videotape ratings of patients manifesting insight during sessions. Future research directions and clinical implications of the SWAP insight scale are discussed.

Key Words: Insight, SWAP, SCORS, CDPS, psychodynamic psychotherapy.

(*J Nerv Ment Dis* 2011;199: 354–359)

The ability to understand ourselves and achieve meaningful insight has been an important process in psychological theory and practice. Socrates provided a simple yet meaningful statement when he said, "the unexamined life is not worth living" (Fowler, 1966). The concept of a patient's inherent capacity to engage in a psychotherapeutic process was first clearly expressed by Freud (1905) in his discussion of analyzability. Freud (1905) recognized early on that certain patient capacities were important factors in fostering various forms of psychological growth, impacting both the treatment and outcome (p. 108). Insight has since been defined as the level of complexity regarding knowledge of the psychological processes of the self and/or others, considered on a continuum from low to high. Some also consider insight as a measure of one's ability to understand the mental states of others and do not include the awareness of the self in their definition (Conte et al., 1990; Piper et al., 1994a). In addition, Fonagy et al. (1991) defined the aspects of insight in his concept of reflective self-functioning as an awareness of the mental states of the self as well as of others. More specifically, insight is described as the progressive ability to understand that human behavior is organized by thoughts, feelings, wishes, and beliefs; each in relation to the self and others (Fonagy et al., 2003; Fonagy and Target, 2003). Most recently, Hill and Castonguay (2007) defined insight in their review as the new event of obtaining such understanding or the tendency to be able to achieve such understanding of the connection between the past and present experiences or between thoughts, feelings, desires, or behaviors (p. 144). These authors also note that the capacity to both use and achieve insight is seen as an essential component of change in most psychotherapies today (Hill and Castonguay, 2007)

Many studies have successfully operationalized insight in an effort to understand its development, causes, and effects on the therapeutic process. Luborsky and colleagues (Luborsky et al., 1988) examined the patient's core conflictual relationship themes, collected by independent raters, to evaluate the expression of self-understanding of patients in session. The results indicated that patient's level of understanding did significantly correlate with treatment outcome. Hoglend et al. (1994) used expert ratings of insight based on audiotape clinical interviews. This study found that the insight levels at pretreatment were positively related to treatment length, the pursuit of additional treatment, and attrition. This study also found that in the course of 2 years, changes in insight were also positively related to other outcomes. In addition, Gelso et al. (1997) found that high levels of transference coupled with high levels of emotional insight predicted a better treatment outcome for 33 clients. Additional studies have also demonstrated that insight was found to be positively correlated to outcome, as assessed by self-report or clinician-rated measures (Piper et al., 1994a, 1994b; Conte et al., 1990). Conversely, Awad (1993) demonstrated that low pretreatment insight was associated with poor psychological functioning at the end of treatment. Based on these findings, perhaps a patient's capacity for insight allows them to more fully engage in treatment. A few studies have also examined patient insight is an outcome criterion that increases in the course of treatment. One study found that self-awareness levels increased in the course of psychotherapy (Vargas, 1954). Grand et al. (2003) also revealed that patients who gained a significant amount of insight throughout treatment were more adaptive to problem solving after treatment termination. These studies support the role of insight in the psychotherapeutic process and development.

Several research measures have incorporated items assessing insight within their scale descriptions. The Vanderbilt Psychotherapy Process Scale (O'Malley et al., 1983) and the Patient Collaboration Scale (Allen et al., 1984) are two scales that contain items regarding the insight construct. Some studies directly assess insight, whereas others describe key qualities of insight without specifically labeling it as such. Therefore, a measure specifically developed to evaluate insight would be beneficial. One scale that includes items specifically for the assessment of insight is the Capacity for Dynamic Process Scale (CDPS; Thackery et al., 1993), which includes nine dynamically oriented affect, insight, and patient capacity specific to object relations items based on the suitability criteria for time-limited dynamic psychotherapy by Sifneos (1979). These items include the patient characteristics of a) appearing introspective, b) integrating affects, c) manifesting verbal fluency, d) manifesting insight, e) perceiving affective aspects of problems, f) differentiating affects, g) differentiating interpersonal events, h) offering positive relationship, and i) collaborating therapeutically. In a recent study by Cromer and Hilsenroth (2010), the CDPS insight subscale demonstrated significant positive relationships with a number of patient- and clinician-rated outcome measures.

Derner Institute of Advanced Psychological Studies, Adelphi University, Garden City, NY.

Send reprint requests to Meaghan E. Lehmann, MA, The Derner Institute of Advanced Psychological Studies, Adelphi University, Hy Weinberg Bldg, 158 Cambridge Avenue, Garden City, NY 11530. E-mail: Meaghan.Lehmann@gmail.com.

What is needed beyond these well constructed existing instruments is not so much a scale that measures insight within the therapeutic process but instead a scale that measures the patients' level of insight within their personality structure. This scale should not only be valid and reliable, but it should be based on observable patient functioning as assessed by independent clinician ratings. Independent clinician ratings provide the complexity necessary to accurately assess level of insight that may be limited in responses to self-report measures on the insight construct. In addition, given the range of theoretical perspectives that provide a definition of insight, it would be useful to provide an assessment of patient insight from a descriptive and experience-near perspective. The Shedler-Westen Assessment Procedure Q-Sort (SWAP-200) is one such measure and is widely used as a diagnostic instrument in personality assessment through observable behaviors and detailed patient accounts of interpersonal interactions. Not only is it clinically derived and empirically based (Shedler and Westen, 1999a; 1999b, 2004; 2007), it also has excellent retest reliability, as well as good interrater, discriminant, and convergent validities with a wide range of external criterion measures (Diener and Hilsenroth, 2004; Shedler and Westen, 2004; Westen and Shedler, 2007; Marin-Avellan et al., 2005; Westen and Arkowitz-Westen, 1998; Westen and Mudderrisoglu, 2003; Shedler and Westen, 1999a; Westen, 1997; Westen et al., 2003; Westen and Weinberger, 2004).

In the present study, we used items from the SWAP-200 to develop an insight scale and examined the construct validity of this index using independent clinical videotape ratings of patient in-session behavior demonstrating both introspection and insight. We also examined the relationship between the SWAP insight scale and the independent clinical videotape ratings of patient complexity of representations as well as their understanding of social causality.

METHODS

Participants

The participants in this study were representative of those actually seeking outpatient treatment in a university-based community clinic. The cases were assigned to treatment and clinicians in an ecologically valid manner based on real world issues regarding aspects of clinician availability, caseload, and others. Moreover, patients were accepted into treatment regardless of disorder or comorbidity. The participants used in this study were 105 patients receiving psychotherapy at a university-based outpatient treatment clinic. As can be observed in Table 1, the sample was predominantly female (68%) and single (63%). The age range was 18 to 65 years, and the mean age for the sample was 29 years (SD, 11 years). This sample consisted of primarily mood-disordered patients with additional relational problems manifested in either axis II or the subclinical traits/features of axis II. Of the participants who had an axis II diagnosis or subclinical traits, 54% fell into cluster B, 31% into cluster C, and 15% into cluster A. The severity of pathology was in the mild to moderate range, which is consistent with what one would expect from a sample of a university-based, community outpatient clinic. Finally, each participant provided written informed consent to be included in this research.

Procedures

Advanced graduate students (13 men and 13 women) in an American Psychological Association–approved clinical PhD program conducted the assessment, feedback sessions, treatment and rated the clinical measures. Each clinician completed graduate course training in descriptive psychopathology and was supervised by a licensed PhD clinical psychologist for a minimum of 3.5 hours per week (1.5 hours individually and 2 hours in a group treatment team meeting). The clinicians completed structured training on the clinical rating scales before rating their patients. They used a number of

TABLE 1. Demographics of Clinical Sample (N = 105)

Age, mean (SD), yrs	29 (11)
Sex, n (%)	
Male	34 (32)
Female	71 (68)
Relationship status, n (%)	
Single	66 (63)
Married	22 (21)
Divorced	16 (15)
Axis I diagnosis, n (%)	
Mood disorder	57 (54)
Anxiety disorder	13 (12)
Eating disorder	2 (2)
Adjustment disorder	14 (13)
V-code	18 (17)
Comorbid axis II diagnosis, n (%)	58 (55)
Subclinical axis II traits/features, n (%)	23 (22)

assessment tools after meeting with the patients for a semistructured clinical interview lasting approximately 2 hours and an interpretive feedback interview lasting approximately 1.5 hours. The clinical interview centered on presenting problems, past history, and relational episodes. The evaluation followed a therapeutic model of assessment (TMA; Hilsenroth, Peters and Ackerman, 2004; Finn and Tonsager, 1992, 1997), which emphasizes focus on collaboration, alliance building, diagnostic impression, and factors contributing to the maintenance of life problems (Hilsenroth et al., 2000; Hilsenroth, 2007). A TMA attempts to optimize the evaluation phase through its utilization of a multimethod assessment (i.e., interview, self-report, performance tasks, and free-response measures). The TMA used in this study consisted of four steps, including three meetings between the patient and clinician and one patient appointment to complete a battery of self-report measures. The three meetings included a) a semistructured diagnostic interview (approximately 2 hours), b) interview follow-up and administration of clinician-rated measures, and c) a collaborative feedback session (@ 1.5 hours), during which the assessment results were shared and explored. The clinical interview, feedback session, and psychotherapy sessions for each patient were videotaped.

Measures

Shedler-Westen Assessment Procedure–200

Therapists used the SWAP-200 to describe their patients after the therapeutic assessment and the first two therapy sessions (approximately 5 to 6 contact hours). The SWAP-200 Q-Sort is a set of statements that describe different aspects of personality and psychological functioning used by a clinician or an interviewer who has a thorough knowledge of the patient. The cards are sorted into categories, from those that are inapplicable or not descriptive to those that are highly descriptive (Shedler and Westen, 1999a, 1999b). The rater arranges the set of 200 personality descriptions into eight different categories ranging from 0 (irrelevant or inapplicable to this patient) to 7 (highly descriptive of the patient).

The Q-Sort creates a fixed distribution of SWAP-200 items that resembles the right half of a normal distribution. The fixed distribution requires the rater to assign a specified number of items to each score category (8 in pile 7; 10 in pile 6; 12 in pile 5, etc.). This process protects against measurement error and heteroscedasticity because of that fact that all clinicians must use each value the same number of times.

Correlation coefficients are then calculated to assess the match between the characteristics of the particular patient and the 10 empirically derived, aggregate descriptions (clinical prototypes and composite

descriptions). The SWAP-200 provides a description of hypothetical, prototypical patients who illustrate a given personality disorder. These hypothetical patient descriptions are referred to as diagnostic prototypes because they accurately reflect the clinical and theoretical understanding of many practicing clinicians. The description of actual patients in a given diagnostic category is referred to as a composite description (Shedler and Westen, 1999a, 1999b).

Several studies support the reliability and validity of the SWAP-200 in the diagnosis of personality disorders. For example, Shedler and Westen (1998), in a study using an earlier version of the SWAP-200, as well as Shedler and Westen (1999a, 2004, 2007) demonstrated strong evidence for convergent and discriminant validity.

Social Cognition and Object Relations Scale

The Social Cognition and Object Relations Scale (SCORS) by Westen (1995) is a narrative-based, dimensional object relations measure designed to assess a variety of dynamic personality features beyond the overt presentation of the patient. The SCORS consists of eight clinician-rated variables that examine the affective and cognitive aspects of an individual's object relations (Stein et al., 2007; Stein et al., 2009). Each of the eight SCORS variables is scored on a 7-point anchored rating scale where lower scores (*e.g.*, 1 or 2) indicate greater pathology and higher scores (*e.g.*, 6 or 7) indicate greater psychological health. The two SCORS variables used in the current study were the Complexity of Representations, which assesses the richness of a patient's representations of self and others and his or her ability to integrate both positive and negative attributes of the self and others, as well as the Understanding of Social Causality, which assesses how well a patient understands why people do what they do. The Complexity of Representations of People (COM) evaluates internal states and how well the patient is able to see internal states in the self and others when reporting narratives. In addition, the COM assesses the patient's relational boundaries and ability to integrate both positive and negative aspects of self and others. A COM score of 5 would describe someone whose "representations of the self and others are stereotypical or conventional, [who] is able to integrate both good and bad characteristics of self and others, [and who] has awareness of impact on others," whereas a COM score of 7 would demonstrate that the individual is "psychologically minded, has insight into self and others, differentiated and shows considerable complexity." The Understanding of Social Causality (SC) measures an individual's ability to understand and accurately perceive social situations and interactions. An SC score of 5 would describe the individual as tending to "provide straightforward narrative accounts of interpersonal events in which people's actions result from the way they experience or interpret situations." An individual with a score of 7 on the SC would be described as someone who "tends to provide particularly coherent narrative accounts of interpersonal events, and to understand people very well, understands the impact of their behavior on others and others' behavior on them." We chose to examine these specific variables because it is likely that an individual capable of high levels of insight may be able to better assess the environment and role dynamics as well as be willing to explore social interpretations.

Consistent with SWAP-200 ratings, the independent SCORS ratings were based on the patient's level of relational functioning at the time of evaluation (*i.e.*, semistructured interview and feedback) and across the first two sessions of psychotherapy (when available). That is, the SCORS variables were dimensionally scored based on relational episodes and self-statements verbally expressed in the course of the psychological evaluation and the first two sessions of psychotherapy. External raters then independently rated the SCORS variables for each participant immediately after viewing videotapes of the clinical interview, feedback, and the first two psychotherapy sessions (when available). External raters in this study consisted of the same pool of PhD graduate clinicians trained in the SCORS rating system

(none provided video ratings for their own patients) or, in some cases, the study supervisor (a licensed PhD Clinical Psychologist). Further details of the SCORS methodology and procedures used in this study are described more fully elsewhere (Peters et al., 2006). The COM and SC SCORS variables used in the study demonstrated good to excellent levels of interrater reliability (Fleiss, 1981; intraclass correlation coefficient [ICC] [1,2] = 0.69 and 0.76, respectively; Peters et al., 2006).

Capacity for Dynamic Process Scale

An individual's capacity for dynamic process refers to the patient's willingness and ability to collaborate therapeutically in addressing problems of an affective and interpersonal nature (Thackery et al., 1993). The CDPS includes nine dynamically oriented personality variables based on the suitability criteria of Brief Dynamic Psychotherapy by Sifneos (1979). These were organized into three conceptual clusters: affect, insight, and relational capacity. In an attempt to capture this constellation of personality variants, the CDPS was developed to assess nine basic areas of functioning based on theoretical and empirical findings related to engagement in the psychodynamic psychotherapy process. These areas are *a)* appears introspective (AI; *i.e.*, the patient seems interested and makes attempts to explore and understand internal events), *b)* integrates affect (*i.e.*, the patient experiences emotion during the interview and neither defends against nor becomes overwhelmed by the affective experience), *c)* manifests verbal fluency (*i.e.*, the patient demonstrates receptive and expressive vocabulary), *d)* manifests insight (MI; *i.e.*, the patient recognizes and understands salient information about self and others), *e)* perceives affective aspects of problems (*i.e.*, the patient discusses emotional components of problems), *f)* differentiates affect (*i.e.*, the patient effectively discriminates qualitatively different internal feeling states), *g)* differentiates interpersonal events (*i.e.*, the patient effectively discriminates attributes of others, of relationships, or of the self relative to others), *h)* offers positive relationship (*i.e.*, the patient conveys friendly acceptance of the clinician), and *i)* collaborates therapeutically (*i.e.*, the patient shares mutual responsibility with the clinician for working together toward exploration and understanding of problems). Each CDPS item is rated on a 5-point scale (*i.e.*, 1, minimal and 5, maximal) using specific rating criteria and a single, overall score for each patient.

The CDPS ratings were based on the patient's level of relational functioning at the time of evaluation (*i.e.*, semistructured interview and feedback), based on relational episodes and self-statements verbally expressed in the course of the psychological evaluation. Consistent with the SCORS ratings, external raters then independently rated the CDPS variables for each participant immediately after viewing videotapes of the clinical interview and feedback. External raters in this study consisted of the same pool of PhD graduate clinicians trained in the CDPS rating system (none provided video ratings for their own patients) or, in some cases, the study supervisor (a licensed PhD clinical psychologist). Further details of the CDPS methodology and procedures used in this study are described more fully elsewhere (Cromer and Hilsenroth, 2010). All CDPS variables used in the study demonstrated good to excellent levels of interrater reliability (Fleiss, 1981; ICC [1, 2] > 0.60; Cromer and Hilsenroth, 2010).

RESULTS

Principal Components Analysis

The 200 SWAP items were individually reviewed in several iterative steps to identify those items that might be most related to patient insight by a team of clinical PhD graduate students and clinical psychologists with extensive experience in using the SWAP. Thirteen items were eventually identified as potentially useful in their ability to assess insight. We then undertook a Principal Components

Analysis factor model with an orthogonal rotation to evaluate and organize these items under consideration (Fabrigar et al., 1999; Floyd and Widaman, 1995). The number of factors retained was determined through an inspection of eigenvalues, percentage variance, and screen plot. Bryant and Yarnold (1995) state that "variables with factor-loading coefficients of at least 0.30 in absolute value 'as loading on the eigenvector' and thus as worthy of consideration in the interpretation of the meaning of the eigenvector (p.106)." Factor analysis of these 13 items revealed a two-factor model solution, accounted for 41% of the variance and included both positive and negative item clusters. Eight items demonstrated factor scores greater than 0.40 on these two factors and there were no dual or multiple loadings. In consultation with the SWAP authors (Shedler, PhD, personal communication (e-mail), April 2008; Westen, PhD, personal communication (e-mail), April 2008), it was observed that two of these eight items were quite specific to borderline pathology and were not included so as not to confound the assessment of insight with this personality constellation. The following two items were removed from the scale: item 79, "tends to see certain others as 'all bad' and loses the capacity to perceive any positive qualities the person may have" and item 162, "expresses contradictory feelings or beliefs without being disturbed by the inconsistency, has little need to reconcile or resolve contradictory ideas." The remaining six items included four items related to positive/presence of insight in patients, whereas the other two related to negative/absence of insight in patients. In Table 2, we list the items included in the SWAP insight scale, two of which have "R" notations, indicating that the item is reverse scored. We then used the original Q-Sort values (0 to 7) of these six SWAP items (two were reverse scored: $7-x$) for each patient to calculate an insight scale total score (simple addition). In addition, the alpha coefficient for the SWAP insight scale demonstrated excellent internal consistency at 0.78.

Convergent Validity

To assess how well the SWAP insight scale evaluates patient levels of insight, we conducted pairwise correlations to examine the construct validity of this index. To do so, we examined the relationship of our scale with the CDPS total score, as well as with the two items that are specific to patient insight, AI and MI. Table 3 illustrates the correlations between the SWAP insight scale and the independent ratings of the CDPS total score, as well as the AI and MI variables. As can be seen from the table, the SWAP insight scale (all six items) was significantly/positively correlated with external ratings of the CDPS

TABLE 3. Convergent Validity of the SWAP Insight Scale and Independent Ratings of the CDPS ($N = 105$)

	CDPS Appears Introspective	CDPS Manifests Insight	CDPS Total
Insight subscale, six items	$r = 0.40$ $p < 0.0001$	$r = 0.45$ $p < 0.0001$	$r = 0.58$ $p < 0.0001$
75-R	$r = 0.42$ $p < 0.0001$	$r = 0.40$ $p < 0.0001$	$r = 0.44$ $p < 0.0001$
82	$r = 0.26$ $p = 0.01$	$r = 0.26$ $p = 0.01$	$r = 0.40$ $p < 0.0001$
89	$r = 0.04$ $p = 0.68$	$r = 0.22$ $p = 0.02$	$r = 0.27$ $p = 0.01$
111	$r = 0.17$ $p = 0.08$	$r = 0.19$ $p = 0.06$	$r = 0.36$ $p = 0.0002$
148-R	$r = 0.37$ $p = 0.0001$	$r = 0.42$ $p < 0.0001$	$r = 0.46$ $p < 0.0001$
183	$r = 0.35$ $p = 0.0002$	$r = 0.38$ $p < 0.0001$	$r = 0.47$ $p < 0.0001$

$df = 103$.
SWAP indicates Shedler-Westen Assessment Procedure; CDPS, Capacity for Dynamic Process Scale.

appears introspective and manifests insight items. Moderate to larger effects are represented in the insight scale and items CDPS AI ($r = 0.40$, $p < 0.0001$) CDPS MI ($r = 0.45$, $p < 0.0001$) and CDPS Total ($r = 0.58$, $p < 0.0001$). Almost all individual items showed significant relationships to at least one or all CDPS items.

In addition to the CDPS, we examined the relationship of the SWAP insight scale with the SCORS variables COM and SC. We can see from Table 4 that the SWAP insight scale overall has convergent validity when related to the SCORS variables. The SWAP insight scale (all six items) was significant/positively correlated with independent ratings of the SCORS COM, and SC variables. Moderate to large effects (Cohen, 1988) are represented with the insight scale and items COM ($r = 0.42$, $p < 0.0001$) and SC ($r = 0.33$, $p < 0.0006$). All individual

TABLE 2. SWAP Insight Scale Items

Item	Description
75-R	Tends to think in concrete terms and interpret things in overly literal ways; has limited ability to appreciate metaphor, analogy, or nuance.
82	Is capable of hearing information that is emotionally threatening (*i.e.*, challenges cherished beliefs, perceptions, and self-perceptions) and can use and benefit from it.
89	Appears to have come to terms with painful experiences from the past; has found meaning in and has grown from such experiences.
111	Has the capacity to recognize alternative viewpoints, even in matters that stir up strong feelings.
148-R	Has little psychological insight into own motives, behavior, and others; is unable to consider alternate interpretations of his/her experiences.
183	Is psychologically insightful; is able to understand self and others in subtle and sophisticated ways.

SWAP indicates Shedler-Westen Assessment Procedure.

TABLE 4. Convergent Validity of the SWAP Insight Scale and Independent Ratings of the SCORS ($N = 105$)

	SCORS: Complexity	SCORS: Social Causality
Insight subscale, six items	$r = 0.42$ $p < 0.0001$	$r = 0.33$ $p = 0.0006$
75-R	$r = 0.24$ $p = 0.02$	$r = 0.13$ $p = 0.18$
82	$r = 0.27$ $p = 0.01$	$r = 0.21$ $p = 0.03$
89	$r = 0.23$ $p = 0.02$	$r = 0.16$ $p = 0.10$
111	$r = 0.28$ $p = 0.004$	$r = 0.25$ $p = 0.01$
148-R	$r = 0.32$ $p = 0.0008$	$r = 0.18$ $p = 0.06$
183	$r = 0.42$ $p < 0.0001$	$r = 0.40$ $p < 0.0001$

$df = 103$.
SWAP indicates Shedler-Westen Assessment Procedure; SCORS, Social Cognition and Object Relations Scale.

items showed significant relationships with at least one or both SCORS variables.

Partial Correlations

To make sure that the SWAP insight scale measures insight specifically and is not simply a measure of general patient functioning as measured by the SWAP, we conducted a partial correlation between the SWAP insight scale and CDPS variable manifesting insight controlling for the effects of the SWAP high functioning scale. We specifically chose the MI variable because of its criterion validity for patient functioning and observable behavior in the clinical interaction. As reported earlier, without controlling for level of functioning (SWAP high functioning), the correlation between SWAP insight scale and CDPS MI was significant ($r = 0.45$, $p < 0.0001$). When controlling for level of functioning (SWAP high functioning) the correlation between SWAP insight scale and CDPS MI remains significant ($r = 0.37$, $p = 0.0001$).

DISCUSSION

The present study developed a scale to evaluate psychological insight as manifested in applied clinical settings using the existing Q-Sort items from the SWAP-200. We then assessed the convergent validity of this SWAP insight scale with established measures of in-session relational functioning and thematic narratives (CDPS and SCORS). The SWAP insight scale demonstrated convergent validity with the SCORS, an object relations measure of personality features, and the CDPS, which assesses a patient's level of engagement in psychotherapy. Therefore, the results demonstrate that these measures, although different, have some overlap on this similar construct. Insight scale scores were significantly related to the CDPS total scores, as well as with AI, wherein the patient demonstrates interest in internal events and attempts to explore and understand them. Insight scale scores were also significantly related to the CDPS variable MI, where the patient demonstrates an ability to recognize and understand salient information about the self and others. Patients with higher scores on the insight scale also had more adaptive levels of manifesting insight and appeared to demonstrate great introspection during clinical interviews. Each of these characteristics support past definitions of insight that encompass reflective self-functioning (Fonagy et al., 1991), as well as the ability to understand one's thoughts, feelings, and beliefs about the self and others (Hill and Castonguay, 2007).

Similarly, higher insight scale scores were reflective of patients with more complex representations of self and others as well as of the ability to integrate both positive and negative attributes of each (SCORS-COM). The COM variable evaluates internal states and how well the patient is able to see internal states in the self and others when reporting narratives as well as assessing the patient's relational boundaries and ability to integrate. Higher insight scale scores were also reflective of patients with greater ability to understand and accurately perceive social situations and interactions (SCORS-SC). An individual capable of high levels of insight may be able to better assess the environment, the role dynamics in that environment and be willing to explore social interpretations. In addition, this ability may also contribute to the facilitating growth within the patient-therapist interaction by allowing more clear and open dialogue about what happens during the therapeutic session. These findings are especially significant because a patient with high levels of insight (including a willingness to introspect on painful dynamics) can perhaps more optimally facilitate and contribute to the process of working through these more difficult therapeutic experiences, thus contributing to treatment outcome. As previous research has found, the insight that emerges from the patient-therapist interaction (Hirsch, 1998) or is achieved throughout the therapeutic process (Luborsky et al., 1988) significantly allows for an opportunity to explore the current relationship with the therapist as a template for interactions outside the therapy room.

In addition, we sought to examine the potential confound of general level of functioning on the relationship among the primary insight variable examined in the study. Therefore, we evaluated the relationship between the insight scale score with the independent video-rated MI variable, partialling out the effects of the SWAP high functioning scale. Partial correlations demonstrated that, even when the effects of level of functioning are controlled for, insight scale scores maintained a significant relationship with videotape ratings of patients manifesting insight during clinical sessions. Therefore, the SWAP insight scale maintains its significant relationship with independent clinical ratings of insight, unrelated to global levels of psychopathology. This is significant because the insight scale upholds specific validity for insight behaviors regardless of the patients' level of functioning. This finding supports previous research conducted by Awad (1993), which demonstrated that low pretreatment insight was associated with poor psychological functioning at the end of treatment. Our study partials out the impact of patient functioning and its relationship with insight, which allows for a more specific assessment of this construct.

Past research that included aspects of insight assessment have used independent clinical ratings (Hoglend et al., 1994; Luborsky et al., 1988; Piper et al., 1994a, 1994b). This study contributes to this existing literature in several ways. First, the SWAP insight scale is unique because it was developed through clinical experience and serves as a clinician-rated scale. Second, this insight scale was correlated with independent videotape rating of actual behavior of a patient in session. This is notable because an independently rated expression of insight via videotape offers data on actual session behavior and helps to establish the criterion validity for our scale. Third, this relationship was independent of global level of patient pathology and supports the specificity of the scale in our understanding of the insight construct. One limitation of note is the mild to moderate range of pathology in the present sample, although these pathology levels are consistent with those observed in many university-based community outpatient samples. Clearly, further research is needed before applying these findings to more severe patient populations. Alternative populations should be assessed to broaden the SWAP insight scale application and empirical validity. Future process and outcome research should continue to identify therapeutic interventions that may effectively promote insight. Aspects of discriminant validity should be assessed, and the measures value should be demonstrated in tests of predictive validity.

The SWAP insight scale is a reliable measure of insight because and maintains its significance even when we controlled for patient level of functioning. Therefore, we see that our construct measures a patient's capacity for insight beyond global levels of functioning. This insight scale may help to inform prognosis, provide an additional treatment-related dimension of interest, might be used to assess progress and outcomes across treatment and further the connection between science and practice by providing empirical data for previously recognized theoretical constructs of importance. With a more reliable and valid understanding of patient pretherapy capacities such as insight, treatment formulation, collaboration, and planning may be enhanced across different psychotherapies. Based on the relationship between in-session interactions and behavior, this scale can provide clinical feedback pertaining to patient characteristics thought to be salient in insight-oriented approaches to psychotherapy. Furthermore, an assessment of this ability may also be optimally used in mentalization, mindfulness, and cognitive-focused treatments.

REFERENCES

Allen JG, Newsom GE, Gabbard GO, Coyne L (1984) Scales to assess the therapeutic alliance from a psychoanalytic perspective. *Bull Menninger Clin.* 48:383–399.

Awad GA (1993) Subjective response in neuroleptics in schizophrenia. *Schizophr Bull.* 21:609–618.

Bryant FB, Yarnold PR (1995) Principal-components analysis and exploratory and confirmatory factor analysis. In L Grimm, PR Yarnold (Eds), *Reading and understanding multivariate statistics* (pp 99–136) Washington (DC): American Psychological Association.

Cohen J (1988) *Statistical power analysis for the behavioral sciences* (2nd ed). Hillsdale, NJ: Erlbaum.

Conte HR, Plutchik R, Jung BB, Picard S, Karasu TB, Lotterman A (1990) Psychological mindedness as a predictor of psychotherapy outcome: A preliminary report. *Compr Psychiatry.* 31:426–431.

Cromer TD, Hilsenroth MJ (2010) Patient personality and outcome in short-term psychodynamic psychotherapy. *J Nerv Ment Dis.* 198:59–66.

Diener M, Hilsenroth M (2004) Multimethod validity assessment of the SWAP-200 dysphoric Q-Factor. *J Nerv Ment Dis.* 192:479–486.

Fabrigar L, Wegener D, MacCallum R, Strahan E (1999) Evaluating the use of exploratory factor analysis in psychological research. *Psychol Methods.* 4:272–299.

Finn S, Tonsager M (1992) Therapeutic effects of providing MMPI-2 test feedback to college students awaiting therapy. *Psychol Assess.* 4:278–287.

Finn S, Tonsager M (1997) Information-gathering and therapeutic models of assessment: Complementary paradigms. *Psychol Assess.* 19:374–385.

Fleiss J (1981) *Statistical methods for rates and proportions* (2nd ed). New York: Wiley.

Floyd FJ, Widaman KF (1995) Factor analysis in the development and refinement of clinical assessment instruments. *Psychol Assess.* 7:286–299.

Fonagy P, Steele M, Steele H, Moran G, Higgitt A (2003) The capacity for understanding mental states: The reflective self in parent and child and its significance for security of attachment. *Infant Ment Health J.* 12:201–218.

Fonagy P, Target M (2003) Evolution of the interpersonal interpretive function: Clues for effective preventive intervention in early childhood. In S Coates, J Rosenthal, D Schechter (Eds), *September 11: Trauma and human bonds* (pp 99–113). Hillsdale, NJ: Analytic Press.

Fonagy P, Steele H, Moran G, Steele M, Higgitt A (1991) The capacity for understanding mental states: the reflective self in parent and child and its significance for security of attachment. *Infant Ment Health J.* 13:200–217.

Fowler HN (1966) *Plato in twelve volumes.* (Vol. 1, p. 38a) Cambridge, MA: Harvard University Press.

Freud S (1905 [1901]) *Fragment of an analysis of a case of hysteria.* standardized ed, 7:3 pp.122.

Gelso CJ, Kivlighan DM, Wine B, Jones A, Friedman SC (1997) Transference, insight, and the course of time-limited therapy. *J Counsel Psychol.* 44:209–217.

Grand T, Rudolf G, Oberbracht C, Pauli-Magnus C (2003) Progressive changes in patients' lives after psychotherapy: Which treatment effects support them? *Psychother Res.* 13:43–58.

Hill C, Castonguay L (2007) *Insight in Psychotherapy.* Washington, DC: American Psychological Association.

Hilsenroth M (2007) A programmatic study of short-term psychodynamic psychotherapy: Assessment, process, outcome and training. *Psychother Res.* 17:31–45.

Hilsenroth M, Peters E, Ackerman S (2004) The development of the therapeutic alliance during psychological assessment: Patient and therapist perspectives across treatment. *J Pers Assess.* 83:332–344.

Hilsenroth MJ, Ackerman SJ, Blagys MD, Baumann BD, Baity MR, Smith SR, Price JL, Smith CL, Heindselman TL, Mount MK, Holdwick DJ Jr (2000) Reliability and validity of DSM-IV Axis V. *Am J Psychiatry.* 157:1858–1863.

Hirsch I (1998) The concept of enactment and theoretical convergence. *Psychoanal Dialogues.* 4:171–192.

Hoglend P, Engelstad V, Sorbye O, Heyerdahl O, Amlo S (1994) The role of insight in exploratory psychodynamic psychotherapy. *Br J Med Psychol.* 67:305–317.

Luborsky L, Crits-Christoph P, Mintz J, Auerbach A (1988) *Who will benefit from psychotherapy?* New York: Basic Books.

Marin-Avellan L, McGauley G, Campbell C, Fonagy P (2005) Using the SWAP-200 in a personality-disordered forensic population: Is it valid, reliable and useful? *J Crim Behav Ment Health.* 15:28–45.

O'Malley SS, Suh CS, Strupp HH (1983) The Vanderbuilt Psychotherapy Process Scale: A report of the scale development and a process-outcome study. *J Consult Clin Psychol.* 51:581–586.

Peters E, Hilsenroth M, Eudell-Simmons E, Blagys M, Handler L (2006) Reliability and validity of the Social Cognition and Object Relations Scale in clinical use. *Psychother Res.* 16:617–626.

Piper WE, Joyce AS, Rosie JS, Azim HS (1994a) Psychological mindedness, work, and outcome in day treatment. *Int J Group Psychother.* 44:291–311.

Piper WE, Joyce AS, Azim HS, Rosie JS (1994b) Patient characteristics and success in day treatment. *J Nerv Ment Dis.* 182:381–386.

Shedler J, Westen D (2007) The Shedler-Westen Assessment Procedure (SWAP): Making personality diagnosis clinically meaningful. *J Pers Assess.* 89:41–55.

Shedler J, Westen D (2004) Refining personality disorder diagnosis: Integrating science and practice. *Am J Psychiatry.* 161:1350–1365.

Shedler J, Westen D (1999a) Revising and assessing Axis II, Part 1: Developing a clinically and empirically valid assessment method. *Am J Psychiatry.* 156:258–272.

Shedler J, Westen D (1999b) Revising and assessing Axis II, Part 2: Toward an empirically based and clinically useful classification of personality disorders. *Am J Psychiatry.* 156:273–285.

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

Sifneos P (1979) *Short-term dynamic psychotherapy.* New York: Plenum.

Stein M, Pinsker-Aspen J, Hilsenroth M (2007) Borderline pathology and the Personality Assessment Inventory (PAI): An evaluation of criterion and concurrent validity. *J Pers Assess.* 88:81–89.

Stein M, Hilsenroth M, Pinsker-Aspen J, Primavera L (2009) Validity of DSM-IV Axis V Global Assessment of Relational Functioning Scale: A multimodal assessment. *J Nerv Ment Dis.* 197:50–55.

Thackery M, Butler SF, Strupp HH (1993) The Capacity for Dynamic Process Scale (CDPS). In ML Canfield, JE Canfield (Eds), *A collection of psychological scales* (pp 57–63). Bartlesville, OK: Research, evaluation, & statistics.

Vargas MJ (1954) Changes in self-awareness during client-centered therapy. In CR Rogers, RF Dymond (Eds), *Psychotherapy and personality change* (pp 145–166). Chicago: University of Chicago Press.

Westen D (1997) Divergences between clinical and research methods for assessing personality disorders: Implications for research and the evolution of Axis II. *Am J Psychiatry.* 154:895–903.

Westen D, Arkowitz-Westen L (1998) Limitations of axis II in diagnosing personality pathology in clinical practice. *Am J Psychiatry.* 155:1767–1771.

Westen D, Muderrisoglu S (2003) Reliability and validity of personality disorder assessment using a systematic clinical interview: Evaluating an alternative to structured interviews. *J Pers Disord.* 17:350–368.

Westen D, Shedler J (2007) Personality diagnosis with the Shedler-Westen Assessment Procedure (SWAP): Integrating clinical and statistical measurement and prediction. *J Abnorm Psychol.* 4:810–822.

Westen D, Shedler J, Durrett C, Glass S, Martens A (2003) Personality diagnosis in adolescence: DSM-IV axis II diagnoses and an empirically derived alternative. *Am J Psychiatry.* 5:952–966.

Westen D, Weinberger J (2004) When clinical description becomes statistical prediction. *Am Psychol.* 59:595–613.

Westen D (1995) Social Cognition and Object Relations Scale: Q-Sort for Projective Stories (SCORS-Q). Unpublished Manuscript.