Published in final edited form as:
J Pers Assess. 2015 ; 97(1): 3–12. doi:10.1080/00223891.2014.930744.

Clinical validity of a dimensional assessment of self- and interpersonal functioning in adolescent inpatients

Greg Haggerty, Ph.D.,
Von Tauber Institute for Global Psychiatry/ Nassau University Medical Center

Mark Blanchard, B.S.,
Von Tauber Institute for Global Psychiatry/ Nassau University Medical Center

Matthew R. Baity, Ph.D.,
Alliant International University

Jared A Defife, Ph.D.,
Emory University

Michelle B. Stein, Ph.D.,
Massachusetts General Hospital/ Harvard Medical School

Caleb J. Siefert, Ph.D.,
University of Michigan-Dearborn

Samuel J. Sinclair, Ph.D., and
Massachusetts General Hospital/ Harvard Medical School

Jennifer Zodan, Ph.D.
Von Tauber Institute for Global Psychiatry/ Nassau University Medical Center

Abstract

The Social Cognition and Object Relations Scale-Global Rating Method (SCORS-G) is a clinical rating system assessing eight domains of self and interpersonal relational experience which can be applied to narrative response data (e.g., Thematic Apperception Test [TAT; Murray, 1943]; early memories narratives) or oral data (e.g., psychotherapy narratives, Relationship Anecdotal Paradigms). In the current study, seventy-two psychiatrically hospitalized adolescents consented and were rated by their individual and group therapist using the SCORS-G. Clinicians also rated therapy engagement, personality functioning, quality of peer relationships, school functioning, global assessment of functioning (GAF), history of eating disordered behavior and history of nonsuicidal self-injury. SCORS-G composite ratings achieved an acceptable level of inter-rater reliability and were associated with theoretically predicted variables (e.g., engagement in therapy; history of nonsuicidal self-injury). SCORS-G ratings also incrementally improved the prediction of therapy engagement and global functioning beyond what was accounted for by GAF scores. This study further demonstrates the clinical utility of the SCORS-G with adolescents.

Personality pathology is an understudied area within the field of adolescent psychiatry. Personality disorders (PD) are common, relatively stable and enduring maladaptive patterns of behavior, thought, impulsivity, or affect. Adolescence is often described as a time of "storm and stress" (Hall, 1904), thus clinicians practice caution when diagnosing adolescents with PDs. Nonetheless, some adolescents do seem to fit PD criteria, experiencing difficulties to a degree that is non-normative during adolescence and which impairs functioning (Shiner & Allen, 2013).

A recent study suggests that rates of PDs among adolescent samples are at least as high, and sometimes higher, than in adults (Shiner & Tackett, in press). Further, personality pathology in adolescents shows rank-order stability across time at a level comparable to with adults (Cohen Crawford, Johnson, & Kasen, 2005; Shiner, 2009). Individuals who eventually evidence personality pathology in adulthood must show at least some problematic patterns during adolescence (American Psychiatric Association, 2013; Shiner & Allen, 2013). Personality pathology in adolescents is also strongly associated with concurrent and future psychiatric problems and maladaptive behaviors (e.g., Bernstein, Cohen, Skodol, Bezirganian, & Brook, 1996; Bornovalova, Hicks, Iacono, & McGue, 2009; Caspi, Roberts, & Shiner, 2005; Cohen et al., 2005; Crawford & Cohen, 2008; de Clercq, van Leeuwen, van den Noortgate, de Bolle, & de Fruyt, 2009; Ferguson, 2010; Johnson et al., 1999; Shiner, 2009; Westen, Betan, & DeFife, 2011).

Psychoanalytic theorists conceptualized PDs as conditions involving dysfunctional self and/or other representations that impaired self and/or interpersonal functioning (Blatt, 2008; Kernberg 1984, 2006). Today, many theories of personality focus on self and interpersonal functioning, such as cognitive-behavioral theories (e.g., Beck, 1999; Linehan, 1993; Young, 1990), trait theories (e.g., Cloninger, 1998; Livesley, Lang, & Vernon, 2003), and interpersonal theories (e.g., Bender & Skodol, 2007; Benjamin, 1996, 2003). In fact, the Diagnostic and Statistical Manual of Mental Disorders 5th edition (DSM-5, APA, 2013) Personality and Personality Disorders Work Group advised reconceptualizing PDs as a "moderate or greater impairment in personality (self/interpersonal) functioning." Findings have shown a link between personality disorders and distorted thinking about self and others (Skodol et al., 2011). This recommendation was based on empirical findings suggesting that maladaptive patterns of self and other representations are useful for conceptualizing personality pathology (Bender et al., 2003; Blatt & Lerner, 1983; Donegan et al., 2003; Wagner & Linahan, 1999; Westen et al., 1990; Zeeck, Hartmann, & Orlinsky, 2006). In short, it is now common to conceptualize personality for adults in terms of self and other representations.

There may also be advantages to examining adolescent personality via the lens of self- and other representations. Studying self and interpersonal functioning in adolescents is especially important given its connection with personality pathology (Bender, Morey, & Skodol, 2011; Livesley, 2007; Morey et al., 2011; Skodol, 2012; Skodol et al., 2011; Tackett, Balsis, Oltmanns, & Krueger, 2009). Impairments in self and interpersonal functioning predict PD outcomes in both adolescents (Defife, Goldberg, & Westen, 2013) and adults (Hopwood et al., 2011).

Researchers (Bornstein, 2011; Huprich, Bornstein, & Schmitt, 2011) have noted the importance of using performance-based methods, such as the SCORS-G, for assessing self and interpersonal functioning. Bornstein (2003) reported that over a 10-year period (1990-1999) more than 80% of published studies in leading journals publishing articles on PDs made use of self-report measures to validate PD symptoms. Though valuable in many respects, self-report measures may also be problematic when assessing some personality constructs in adolescents. In some cases, people may be either unwilling or unable to accurately report on personality (Ganellen, 2007; Huprich et al., 2011). This may be especially true of people with personality pathology as insight into problematic behaviors or patterns may be limited (Ganellen, 2007; Shiner & Allen, 2013). Relative to people without personality pathology, those with personality pathology are more likely to experience impairments in accurately gauging their effect on others (Klonsky, Oltmanns, and Turkheimer, 2002). This may limit their ability to accurately report on interpersonal functioning or effectiveness. A person's mood (Huprich et al., 2011) as well as implicit psychological processes (Kihlstrom & Klein, 1997; Kunda & Thagard, 1996; McNamara, 2005; Shevrin & Dickman, 1980; Westen & Gabbard, 2002) may also bias response to self-report inventories. The latter has been demonstrated in research using priming (Bargh, Bond, Lombardi, & Tota, 1986; DeMarree, Wheeler, & Petty, 2005; Hull, Slone, Meteyer, & Matthews, 2002; Markman & McMullen, 2003; Mussweiler, 2003; Wheeler, DeMarree, & Petty, 2005).

Of course, none of the arguments just reviewed are intended to assert that self-report inventories are useless or invalid. In fact, to the contrary, we argue that self-report inventories are essential to any good personality assessment. Their use may, however, be enhanced by comparing self-reported data with data gleaned from other sources in a multi-method assessment (Meyer et al., 2001). In addition to obtaining a strong history, multi-method assessments often collect data with other assessment methods (e.g., performance tests; clinical rating scales). Growing interest in thoroughly assessing adolescent personality highlights the need for strong assessment tools. While many self-report personality inventories for adolescents currently exist, there is a lack of clinician-rated tools for assessing adolescent personality in clinical settings.

The current study examines the reliability and validity of the Social Cognition and Object Relations Scale-Global Version (SCORS-G; Stein, Hilsenroth, Slavin-Mulford, & Pinsker, 2011; Westen, 1995) with an adolescent inpatient sample. The SCORS-G is a clinician-rated dimensional measure of self and interpersonal functioning. It was originally designed by Westen (Westen, 1991; Westen, Lohr, Silk, Gold, & Kerber, 1990) to assess object relations. The SCORS-G demonstrates strong reliability across raters and narrative data sources (Huprich & Greenberg, 2003). It has also been shown to differentiate patients with PDs from patients without PDs, differentiate among different types of PDs, and differentiate levels of dysfunction (for reviews see Ackerman, Clemence, Weatherill, & Hilsenroth, 1999).

Some studies have examined the SCORS-G with adolescent samples (Bambery & Porcerelli, 2006; Porcerelli, Cogan, & Bambery, 2011; Conway, Oster, & McCarthy, 2010; Gramache, Diguer, Laverdiere, & Rousseau, 2012; Gregory & Mustata, 2012). For example, the SCORS-G has been used to identify subtypes of adolescent sexual offenders (Gramache et

al., 2012), and to understand self and interpersonal processes with adolescents who engage in cutting behavior (Gregory & Mustata, 2012).

The most significant study to examine the SCORS-G in an adolescent setting was conducted by Defife, Goldberg, and Westen (2013). They recruited 294 psychologists and psychiatrists who were currently treating an adolescent (13-18 years old) experiencing personality pathology. Each clinician completed a packet of measures on an adolescent patient he or she was treating who was experiencing personality pathology. The packet included the SCORS-G as well as clinical ratings of the adolescent's functioning and history. The results revealed the SCORS-G items showed medium-to-large effect size differences for those diagnosed with personality disorders versus those who were not. Ratings on the SCORS-G composite rating was also significantly positively related to composite ratings of adaptive functioning, school functioning and negative correlated with externalizing behavior, and prior psychiatric history. The SCORS-G composite rating predicted variance in adaptive functioning domains above and beyond the DSM-IV personality disorder diagnosis.

No studies, to our knowledge, have investigated the inter-rater reliability and construct validity of the SCORS-G with adolescent inpatients. Research has yet to investigate whether two different clinicians with their own unique experiences with an adolescent can reliably rate an adolescent using the SCORS-G. Most existing research using the SCORS-G with adolescents includes using the SCORS-G to rate narrative data (i.e., TAT) or video recorded session material where both observers are providing ratings on the same exact material.

The study sought to extend the work of Defife et al. (2013) by examining the validity and reliability of the SCORS-G with a sample of adolescent inpatients. Defife et al. (2013) utilized a predominately outpatient sample and had one clinician complete all ratings for one patient. The current study is distinctive in that it is the first to use ratings from inpatient unit therapists and data from the patient's chart to examine the construct validity of the SCORS-G. Previous studies have found that SCORS-G ratings are positively correlated with measures of adaptive functioning (Defife et al., 2013; Stein, Hilsenroth, Pinsker-Aspen, & Primavera, 2009), school performance, and peer relationships (Defife et al., 2013; Defife & Westen, unpublished data). Self and interpersonal functional impairments have been implicated in adolescents engaging in NSSI (Nock, 2010) and eating disordered behavior (Cross, Westen, & Bradley, 2011).

Based on the research available thus far, the current study has several hypotheses. First, we predict that the SCORS-G will demonstrate good inter-rater reliability. Second, we hypothesize that SCORS-G will demonstrate evidence of construct validity by demonstrating a positive correlation with clinical ratings of: the Global Assessment of Functioning scale (GAF); clinician-rated patient engagement in group and individual psychotherapy; clinician-ratings of patients' chronic personality functioning, school functioning, and peer relationships. Finally, we hypothesize that greater impairment as indicated by the SCORS-G will be related chart reported histories of non-suicidal self-injury and history of eating disordered behavior.

Methods

Procedure

Parents and legal guardians were approached regarding participation in this study by a research team member. Those who provided informed consent were entered into the study. All patients, whether they were enrolled in this study or not, received individual psychotherapy and group therapy as part of their treatment on the unit. Patients were assigned to a therapist on the first business day after admission in an ecologically valid manner based on clinician availability and caseload. The patients' individual and group therapist completed GAF ratings based on the patients' chart data from the emergency department notes, admitting nurses' notes and first unit psychiatrists' notes. The patients' individual and group therapists also completed clinical ratings at discharge which included the SCORS-G, ratings of therapeutic engagement and a clinical data form which contained ratings of different aspects of the patient's life, demographic information, and information about the patient's development. A member of the research team also reviewed the chart for history of nonsuicidal self-harm and eating disordered behavior.

Participants

The sample consisted of 72 patients (45% of the total patients admitted to the unit), 52.8% male, consecutively admitted to the adolescent psychiatric inpatient unit of a large northeastern hospital. This study was approved by the hospital's IRB of record. Patients were between the ages of 13 – 17 with a mean age of 15.7 (SD = 1.18). Ethnic makeup of the sample was as follows: 40.8% Caucasian, 25.4% African American, 25.4% Latino/ Hispanic/Spanish, 5.6% Other and 2.8% Asian. The primary diagnoses for these 72 patients were as follows: 64% Mood Disorders, 30% Conduct Disorder/Oppositional Defiant Disorder, 2% Impulse Control Disorder, 3% Psychosis, 1% Post Traumatic Stress Disorder. We found no significant differences in age, gender, or diagnosis between those who consented to the study versus those who did not consent. Patients who showed cognitive impairment or IQ below 70 were excluded from this study (4 patients). Twenty-three percent of the sample reported a history of nonsuicidal self-mutilation and 4% had a history of eating disordered behavior.

Setting

The adolescent inpatient unit in this study is a 12-bed locked facility at a large northeastern hospital. The unit provides treatment for adolescents in acute distress. The average length of stay for patients on this unit is 10.81 days (*SD* = 5.23). While on the unit, adolescents receive psychopharmalogical treatment as deemed necessary, two or three individual psychotherapy sessions per week, three general group sessions per week, a weekly structured anger management group, a weekly substance use psycho-education group, daily academic programming, and daily recreation therapy. On average, participants received three individual therapy sessions (SD = 1.56) and six group sessions (SD = 3.08) during their stay.

Raters

Participating clinicians included a licensed clinical psychologist with over five years of experience with advanced training in assessment and five advanced clinical psychology doctoral students (interns and externs) who had completed advanced coursework in assessment at an APA-accredited clinical psychology PhD or PsyD program and were supervised by a licensed clinical psychologist. The ratings were evenly distributed between the licensed clinical psychologist, psychology interns and externs.

Rating Procedure

The group therapist assigned to complete assessments for the patient had to be a co-leader of at least three groups each week on the unit. Individual and group psychotherapists completed GAF ratings based on admission data (emergency department records and physical assessment, the first psychiatry attending assessment note and the first nursing assessment note in the chart). The individual and group therapists also completed clinical assessments at the patient's discharge which included the Social Cognition and Object Relationship Scale-Global Rating Method (SCORS-G; Stein, Hilsenroth, Slavin-Mulford & Pinsker, 2011). The individual and group therapists were blind to each others' ratings both for admission and discharge measures, including any GAF scores given by the Emergency Department or unit clinicians. The patient's individual therapist utilized all available information (i.e., chart, direct interaction with patient, feedback from unit staff) to complete the Clinical Data Form-Adolescent Version (CDF-A; Westen, Shedler, Durrett, Glass, & Martens, 2003) which included dimensional ratings of the patient's chronic level of personality, peer relationships, and school functioning. The CDF-A was completed at the patients' discharge from the unit. The therapists in this study were part of the unit clinical staff and, although added to the research protocol, were not aware of any of the study hypotheses. A member of the research staff who was not part of the clinical team reviewed the patient's chart and noted any history of eating disordered behavior and nonsuicidal self-injury when the patient was discharged from the unit.

Reliability Training

The initial training for each rater was a two-hour session led by the first author. Subsequent hour long group training meetings were held biweekly for the duration of the study. The training included a review of training materials for all of the study measures at the beginning of the study. At the biweekly meetings, members would blindly assess a patient they all knew well who was not part of the study and scoring discrepancies would be discussed. The group would contact an expert on the SCORS-G to discuss the issues further to ensure accurate scoring of study measure if scoring questions arose. A total of 32 reliability trainings occurred including the initial training. Not all raters met for all 32 meetings because psychology interns and externs rotated through the unit for six months.

Measures

Clinical Data Form-Adolescent Version (CDF-A; Westen, Shedler, Durrett, Glass, & Martens, 2003)—The CDF-A is used to gather a wide range of demographic and diagnostic information about the patient. For this study we investigated several ratings

on the CDF which include: Chronic level of personality functioning (1 = severe personality disorder to 5 = high functioning); quality of peer relationships (1 = very poor or absent to 5 = very good); and school functioning (1 = failing/dropped out to 5 = working to full potential). These ratings were completed by the patient's individual therapist. All available background information was taken into consideration. The CDF has been used before demonstrating high reliability and validity with independent expert observers (Dutra, Campbell, & Westen, 2004; Westen, Muderrisoglu, Fowler, Shedler, & Koren, 1997). Defife, Drill, Nakash, and Westen (2010) also found that adaptive functioning and developmental history variables measured using the CDF show high degrees of correspondence and agreement between clinician-rated and patient-rated assessments.

Social Cognition and Object Relations Scale – Global Ratings (SCORS-G; Stein, Hilsenroth, Slavin-Mulford, & Pinsker, 2011)—The SCORS-G is a clinician-rated measure of a patient's representations of self and significant others. The scale consists of eight theoretically constructed variables that are scored on a 7-point Likert-type scale, where a lower score indicates more pathological responses and a higher score suggests healthy functioning. The eight variables are as follows: 1) Complexity of Representation (COM) reflects the richness of one's representations of self and others, one's ability to recognize internal states in self and others, and one's ability to integrate both positive and negative aspects of self and others; 2) Affective Quality of Representations (AFF) assesses one's expectations of others (positive or negative) and one's evaluation of past relationships; 3) Emotional Investment in Relationships (EIR) relates to one's capacity for intimacy and emotional sharing; 4) Emotional Investment in Moral Standards (EIM) broadly reflects one's ability to think about moral questions and show genuine compassion towards others; 5) Understanding of Social Causality (CS) assesses the extent to which one understands human behavior, or why people act the way they do in various situations; 6) Experience and Management of Aggressive Impulses (AGG) reflect one's ability to tolerate and appropriately express anger; 7) Self-Esteem (SE) assesses one's self esteem; and 8) Identity and Coherence of Self (ICS) assesses one's level of integration versus fragmentation. The SCORS-G is a modification of the original SCORS which only included increasing the scale ratings from 5 to 7 scale points. Newer and modified scales/dimensions were also included to the SCORS-G based on research findings which included the original four (Complexity/ Cognitive Structure of Representations, Affective Quality of Representations, Understanding of Social Causality, and Capacity for Emotional Investment in Relationships and Morals). Scales/dimensions such as Experience and Management of Aggression, Self-esteem, and Identity and Sense of Self were added to the SCORS-G and the Emotional Investment in Relationships and Morals was broken down into two separate variables. The SCORS has shown good to excellent reliability when used to rate semi-structured interview data, TAT narratives, early memories narratives, dream narratives, and other clinical data such as psychotherapy session material (cf. Stein, Hilsenroth, Slavin-Mulford, & Pinsker, 2011). For this study we averaged all 8 items of the SCORS-G. Past research (Defife et al., 2013) has shown the average SCORS-G to be significantly related to clinician-ratings of adolescent patients. The advantages to using an average score is that instead of having one item for each subscale there are eight items that can be used to measure the construct, object relations. We then calculated a *SCORS-G composite rating* by averaging the individual and

group therapist's mean SCORS-G rating. The SCORS-G ratings were based on the clinician's interactions with the patient both in therapy and on the unit. All available information including background information and behavioral observations of the patient made by the staff were used in the clinicians' SCORS-G ratings. Cronbach alpha for the SCORS-G composite rating which averaged together the two raters composite scores was .87.

Global Assessment of Functioning (GAF; American Psychiatric Association, 2000)—The GAF is a 100-point clinician-rated scale created to assess a patient's overall level of functioning. The scale is behaviorally anchored to help guide clinical ratings. Both the patient's individual and group therapist completed the GAF blind to each others' ratings at admission. Every therapist provided ratings based on all available data about the patients' functioning when admitted and again at discharge. The two initial GAF ratings were averaged together to form a *mean admission GAF score*. Given that raters scores were averaged together, we calculated the Intraclass Correlation Coefficient (ICC; Shrout & Fleiss, 1979) using a one-way random approach to assess reliability ($ICC_{1,2}$). ICCs of < .40 are considered poor, fair = .40 - .59, good = .60-.74, and >.75 is considered excellent (Shrout & Fleiss, 1979). Raters in the present study obtained an ICC ($_{1,2}$) of .58 for the admission GAF which falls into the fair range. Discharge GAF ratings were not used because the ratings' reliability was lower ($ICC_{1,2}$ = .40).

Clinician Therapy Engagement Rating—The patients' individual and group psychotherapists rated the overall level of engagement and participation demonstrated by the patient in their respective treatment modalities during the length of their hospitalization. They rated them on a scale from 1 (“not engaged”) to 5 (“very engaged”). This rating was done blind to the other clinicians' ratings of each patient as well as self-report questionnaire results. These ratings were completed at the patient's discharge. Raters were instructed to rate the patients overall averaged level of engagement. Training for this included reviewing patients known to the raters but not part of the study to provide guidance on scoring this measure. Raters had to evidence at least an ICC of .70 during the training sessions for this measure.

History of Eating Disordered Behavior and Nonsuicidal Self-Injury (NSSI)—A member of the research staff who was not part of the clinical team reviewed the patients' charts for a history of eating disordered behavior (i.e., binging, purging, restricting) and rated it *1* if the patient has no history of eating disordered behavior and *2* if they did. The same rating was used for a presence or absence of nonsuicidal self-harming behavior where *1* was assigned if the patient did not have a history of NSSI and *2* if there was.

Results

Means and standard deviations for all study measures are reported in Table 1. The first aim of the study was to investigate the inter-rater reliability of the SCORS-G in clinical practice with an adolescent inpatient sample. Given that the raters were not necessarily the same across all patients a one-way random effects model was calculated (ICC ($_{1,2}$) = .64; 95% CI [.40-.78]) and found to be in the good range (Shrout & Fleiss, 1979). To improve the clinical

applicability of the SCORS-G, we also calculated the single measure ICC ($_{1,1}$) as most practitioners are unlikely to use a second rater for their patients. This resulted with a reliability estimate in the fair (.49; 95% CI [.25-.64]) range.

To test our remaining hypotheses, we conducted Pearson product moment correlations between the SCORS-G composite ratings and the remaining study variables. Table 2 contains the inter-rater reliabilities (ICC $_{1,2}$) of the SCORS-G composite and subscales as well as correlations between the SCORS-G composite rating and subscale with the criterion variables. The addition of the subscale reliabilities and correlations with criterion variables was for generalizability to past SCORS-G research which has focused on subscale scores. As predicted, we found that the SCORS-G composite rating was significantly positively correlated with the GAF admission mean score, global ratings of engagement in individual and also group psychotherapy, ratings of chronic personality functioning, school functioning, and peer relationships. Results also showed that SCORS-G composite ratings were significantly negatively correlated with history of non-suicidal self-injury and also eating disordered behavior.

Post-Hoc Analyses

We performed some secondary analyses to investigate whether the SCORS-G ratings provide incremental validity above the GAF scores in predicting therapy engagement and functioning. Although the GAF was not designed to predict engagement, research (Greeno, Anderson, Shear, & Mike, 1999; Turner, Boden, & Mulder, 2013) has shown the GAF scores are positively correlated to treatment engagement ratings. We conducted hierarchical linear regressions in which we entered the GAF score in the first block and the SCORS-G composite ratings in the second block. In the first regression, we calculated a *global therapy engagement composite score* to see whether the SCORS-G ratings provided incremental validity for this dependent variable. We calculated the therapy engagement composite score by adding together the engagement ratings for the patients' individual and group therapy. For the second regression, we calculated an *overall functioning composite score* to see whether SCORS-G ratings provided incremental validity. We calculated this composite score by adding together the clinician-ratings of the patients' *chronic level of personality functioning*, *quality of peer relationships* and *school functioning*.[1] In both regressions, SCORS-G rating provided incremental validity above and beyond GAF scores. These regression results are presented in Table 3.

We also ran hierarchical linear regressions for the dependent variables, *global therapy engagement composite score*, *overall functioning composite score*, in which we entered the SCORS-G composite rating in the first block and GAF in the second block to see if GAF scores added any incremental validity to the SCORS-G composite ratings. Our results showed that GAF did not add any incremental validity above and beyond the SCORS-G composite ratings. The results are presented in Tables 4. These results further show the

[1]To support the creation of this composite, we ran correlations between the *overall functioning composite score* and the three ratings that made up this composite and found that all three ratings correlated above .74 with the overall composite rating. We also used the composite score instead of the individual ratings of personality, school and peer functioning to limit the number of analyses. We ran a principle component analyses with the three items that made up the scale and found that one factor accounted for 60% of the variance.

usefulness in the assessment of adolescents' self and interpersonal functioning and demonstrate that this assessment adds additional information to our understanding of these patients.

We also conducted analyses in which ratings completed by the individual therapist were compared to the group therapist ratings rather than comparing averaged together ratings. This was to address issues of shared method variance. We have included these in Table 5. In order to compare our results with Defife et al.'s (2013) data, we requested correlational results (Defife & Westen, unpublished data) from their dataset that matched the analyses we conducted in this study. Table 6 contains the results of a focused effect size comparisons between data from Defife (unpublished data) and our data (Ferguson & Takane, 1989; Field, 2001)[2].

Discussion

The present study is the first, to the authors' knowledge, to investigate the clinical validity of a dimensional measure of self and interpersonal functioning, the SCORS-G, with an adolescent inpatient sample. Much of the past research studies using the SCORS-G rating involved two highly trained individuals rating the same exact data sources (i.e., TAT or early memory narratives, video recorded psychotherapy sessions), whereas this study investigated the SCORS-G reliability when individuals rate different data sources (i.e., their unique experience with the patient). Although this study did train raters, they were clinicians on the unit with primary clinical duties. Thus the reliability protocol had to be shortened and less structured to meet the real world demands of the volunteer raters than in past studies using the SCORS-G. Despite this difference, the current study showed that two clinicians with different experiences with a patient could also provide reliable and valid SCORS-G ratings. Not only does this study extend the utility of the SCORS to a younger patient sample, but it also demonstrates that a dimensional assessment of personality can be implemented in treatment environments with minimal time commitments. Once trained, none of the clinical raters reported feeling burdened by the procedures.

Our results further those found by Defife et al. (2013) and support the clinical validity of the SCORS-G for assessing self and interpersonal functioning with adolescent inpatients. Following Defife et al. (2013), we used the SCORS-G composite rating instead of the scale scores because more items measuring a construct increase the measures' reliability. Even though Defife et al. did not use the same analyses between the items of the CDF-A and SCORS-G, they were able to provide us with the comparison analyses from their dataset. The focused effect size comparisons show that our results are in line with their data. Table 6 contains the results of these analyses. As our hypotheses predicted, higher SCORS-G ratings were positively correlated with clinician related engagement in both group and individual psychotherapy. Stated differently, those who have healthier and more mature ways of viewing themselves and others were more able to engage in psychotherapeutic treatment. Another possibility is that engagement in therapy may have biased the clinical raters to rate

[2]Focus effect size comparisons were conducted using equation 1 from Field, 2001, p. 163 and equation 12.12 from Ferguson and Takane, 1989, p.208.

the patient higher on the SCORS-G because individual therapist was one of the raters for the SCORS-G and also the rater for the individual therapy engagement rating. However, our results from Table 5 show that when looking at group therapist-rated SCORS-G ratings to individual therapist-rated engagement in individual therapy there is still a significant relationship thus minimizing the likelihood that the rater bias was responsible for the relationship between therapy engagement and SCORS-G ratings. Past research with adult samples have also reported similar findings. (Pinsker-Aspen, Stein, & Hilsenroth, 2007).

We also found that the SCORS-G composite ratings were positively correlated to the patients' GAF scores. This provides convergent validity as similar results were found in past SCORS-G research (Stein, Hilsenroth, Pinsker-Aspen, & Primavera, 2009). We would expect that those who have healthier self and interpersonal functioning capacities would lead to better coping capacities and overall functioning. Adolescents who had healthier SCORS-G ratings were rated healthier by their clinicians in terms of personality functioning, school functioning, and peer relationships. This finding helps to underline to relationships between the SCORS-G and GAF scores, as the GAF ratings take into consideration the adolescents' overall chronic level of personality, peer, and school functioning. The fact that the SCORS-G ratings showed incremental validity above the GAF in predicting both therapy engagement and overall functioning suggest the SCORS-G might be a more useful tool in determining functioning on the unit. Correlations between the GAF and *chronic level of personality functioning* ($r = .39$, $p < .01$), *quality of peer relationships* ($r = .17$, *ns*), and *school functioning* ($r = .31$, $p = .01$) indicate these variables are related but not redundant. When we compared the SCORS-G composite ratings to past behavior as reported in the patients' file, the ratings were negatively correlated with histories of nonsuicidal self-injury and eating disordered behavior. Past research findings reveal that people who exhibit these kinds of maladaptive behaviors often demonstrate problems in their self and interpersonal functioning (Whipple & Fowler, 2011).

Our study was also the first to demonstrate the construct validity of the SCORS-G ratings in an adolescent inpatient sample. We should also note that although the clinical raters took part in bi-weekly reliability trainings, this procedure was not as stringent as in previous research using the SCORS-G and as highlighted in the SCORS-G training manual (Stein, Hilsenroth, Slavin-Mulford, & Pinsker, 2011). This may be the reason why the inter-rater reliability of the SCORS-G and GAF were not in the excellent range. We also would note that reliabilities derived from two clinicians with unique and slightly different experiences with a patient may frequently be lower than if two clinicians rated the exact same material. However, future research needs to explore this question. Our results support the idea that the SCORS-G composite score can be used reliably with less training but comprehensive training may be needed to reliably rate patients on some of the items of the SCORS-G that showed lower reliability scores in our sample. Overall, these initial results are supportive of the use of the SCORS-G with a hospitalized adolescent sample.

Several limitations of this study require discussion. We tried to be conservative by using measures in which the ratings completed by the individual and group therapist were averaged together. As a result, some of the clinical ratings were supplied by the same clinician. For example, the individual therapist who completed the global rating of

engagement for individual therapy could also have been one of the two raters who completed the GAF and SCORS-G ratings. This could lead to an interdependency of SCORS-G composite ratings and their correlates and inflate the relationships between variables due to shared method variance (cf. Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). We performed post-hoc analyses comparing ratings completed by individual therapist with ratings completed by the group therapist. Although the results of these analyses demonstrate smaller effect sizes, the relationship between the SCORS-G composite rating and the criterion variables supported the construct validity of the SCORS-G. Future research needs to look at comparing the SCORS-G with ratings completed by separate people who are not clinicians on the unit and are blind to other clinical data. Even so we would like to highlight that our research also incorporates multiple clinical ratings (e.g., patients' group and individual therapist), as well as patients' behavioral histories from the chart which is aggregated through multiple outside sources (e.g., parents, teachers, past clinicians, etc.). Our study utilized a multimodal assessment which has advantages to studies which solely rely on patient self-reports or clinical ratings.

The small sample size for this study was another limitation. Not all of the consented patients received clinical ratings. This was largely due to the fact that clinician raters did not rate patients with whom they did not have adequate exposure. Our sample also only contained a small number of people who had a history of eating disordered behavior and our results related to this should be taken with caution. Adolescent inpatient samples present unique challenges in treatment compliance which impacts the rate at which data can be collected. We hope that the results of this study will spur on future research on the topic with this population using larger samples. Clinical samples are inherently more difficult to obtain and a sample of hospitalized adolescents is especially difficult. As a result, we feel that the clinical importance of using an at-risk hospitalized clinical sample is a strength which somewhat offset the issues related to sample size.

We incorporated the Global Assessment of Functioning into this study because, at the time, it was broadly used in all clinical settings. Additionally, the GAF ratings are routinely used in assessing the suitability of discharge settings of patients. The GAF also gives us a global measure of how the patient is functioning. However, limitations of the GAF are well documented and include how it is difficult to discern whether a change in GAF was due to poorer functioning in school or in their interpersonal relationships (Winters, Collett, & Myers, 2005), and the poorer reliability in clinical (.54-.65; Hall, 1995; Jones, Thornicroft, Coffey, & Dunn, 1995; Loevdahl & Friis, 1996; Michels, Siebel, Freyberger, Stieglitz, Schaub, & Dilling, H. 1996) versus research settings (.86-.90; Hilsenroth, Ackerman, Blagys, Baumann, Baity, Smith, Price, Smith, Heindselman, Mount, & Holdwick, 2000;Soderberg, Tungstrom, & Armelius, 2005; Tracy, Adler, Rotrsen, Edson, & Lavori, 1997). Bacon, Collins, and Plake (2002) found that GAF ratings were more influenced by symptom rather than adaptive functioning. Many clinicians in regular practice do not receive reliability training and often do not consult the DSM for guidance on scoring. The GAF is also poorly anchored which leads to issues in reliability. The SCORS-G, in contrast, is brief, well anchored and assesses individuals on a number of useful aspects of function such as their regulation of aggression and investment in relationships which the GAF does not.

Some strengths of the current study aside from those mentioned previously were that the SCORS-G was used in the way we would hope it would be used in non-research everyday clinical use. Many clinical ratings are only used in research settings and for various reasons not incorporated in use clinically. For example, a treatment team would approach treatment differently for a hospitalized adolescent patient who shows low scores (more pathological) on both the SCORS-G item *Experience and Management of Aggressive Impulses* and also on *Emotional Investment in Values and Moral Standards* versus another patient with low scores on *Experience and Management of Aggressive Impulses* and also *Understanding of Social Causality*. The first patient's aggression may be motivated more by antisocial/psychopathic needs of power and lacking in empathy, while the second patient's aggressive behavior may be more a result of an unsophisticated appreciation of the actions and intentions of others. The treatment with the second patient would be focused on helping to improve interpreting the motives of others and getting the patient to view alternative explanations for others' behavior beyond just aggressive. This treatment tact would be fatally flawed with the first patient. A better treatment plan with the first patient would focus on a clear and fair adherence to the structure and rules of the unit. The goal would be to have the patient understand the natural fair consequences of his or her behavior.

Given that the GAF has not been included in DSM-5, the SCORS-G could continue to provide clinicians with a reliable and valid assessment of self and interpersonal functioning. The DSM-5 Personality Disorders Work Group proposed general definition for personality disorders highlighted “a broad failure to develop important personality structures and capacities needed for adaptive functioning” (Skodol et al., 2011, p. 17). This revised definition depends upon adaptive failures in the domains of self and interpersonal functioning. Although the DSM-5 ultimately did not revise the personality disorders diagnoses from DSM-IV, they did add an alternative model for personality disorders in Section III for further evaluation which highlights in its general criteria for personality disorder that a person must demonstrate a “moderate or greater impairment in personality (self/interpersonal) functioning.” The DSM-5 also included a Level of Personality Functioning Scale (LPFS). The SCORS-G is similar yet different from the LPFS. Both measures touch on similar underlying constructs such as affect regulation, identity and relationship quality. The scoring instructions and purposes for creation are different. The SCORS-G allows for raters to assess each construct separately providing a more nuanced and complex assessment of a person on these separate but related underlying constructs. The LPFS instructs the rater to provide a single rating/assigned level that best assesses the person's current *overall* impairment level of personality functioning. The LPFS was designed to aid in the diagnosis of a personality disorder while the SCORS-G was designed without diagnosis in mind but to broadly assess social cognition and object relational abilities. Currently, the SCORS-G has more empirical data to support its validity and reliability having been researched with clinical samples, and especially with adolescents. Future research should compare the two measures to see how much overlap exists and which is better at predicting what clinically.

In sum, this study demonstrates the clinical utility and validity of the SCORS-G for assessing self and interpersonal functioning with an adolescent inpatient sample. We not

only showed how the SCORS-G ratings were related to other clinical ratings but also demonstrated its relationship to actual reported behavior of the patients which is an extension to the previous research on the topic (Defife et al., 2013). These results also support the use of the SCORS-G in both broadening the research literature of personality pathology in adolescents and providing a useful clinical assessment measure for inpatient clinicians. Assessing self and interpersonal functioning on admitted adolescents could help clinicians better tailor treatments to the patients presenting strengths or deficits and ultimately formulate more focused treatment plans.

Acknowledgments

Funding: The preparation of the manuscript was funded in part by a grant awarded to the first author from the American Psychoanalytic Association and by National Institute of Mental Health (NIMH) Grant No. 1R21MH097781-01A1 received by the first author.

References

Ackerman SJ, Clemence AJ, Weatherill R, Hilsenroth MJ. Use of the TAT in assessment of DSM-IV cluster B personality disorders. Journal of Personality Assessment. 1999; 73(3):422–448. [PubMed: 10689653]

American Psychiatric Association. DSM 5. American Psychiatric Association; 2013.

Bacon SF, Collins MJ, Plake EV. Does the Global Assessment of Functioning (GAF) assess functioning? Journal of Mental Health Counseling. 2002; 24:202–212.

Bambery M, Porcerelli JM. Psychodynamic therapy for oppositional defiant disorder: Changes in personality, object relations, and adaptive function after six months of treatment. Journal of the American Psychoanalytic Association. 2006; 54(4):1334–1339. [PubMed: 17354508]

Bargh JA, Bond RN, Lombardi WJ, Tota ME. The additive nature of chronic and temporary sources of construct accessibility. Journal of Personality and Social Psychology. 1986; 50:869–878.

Beck, A. Cognitive aspects of personality disorders and their relation to syndromal disorders: A psychoevolutionary approach. In: Cloninger, CR., editor. Personality and psychopathology. Washington, DC: American Psychiatric Association; 1999. p. 411-429.

Bender DS, Farber BA, Sanislow CA, Dyck IR, Geller JD, Skodol AE. Representations of therapists by patients with personality disorders. American Journal of Psychotherapy. 2003; 57:219–236. [PubMed: 12817552]

Bender DS, Morey LC, Skodol AE. Toward a model for assessing level of personality functioning in DSM-5, Part I: A review of theory and methods. Journal of Personality Assessment. 2011; 93:332–346. [PubMed: 22804672]

Bender DS, Skodol AE. Borderline personality as a self-other representational disturbance. Journal of Personality Disorders. 2007; 21:500–517. [PubMed: 17953503]

Benjamin, LS. Interpersonal diagnosis and treatment of personality disorders. New York: Guilford Press; 1996.

Benjamin, JS. Interpersonal reconstructive therapy: Promoting change in nonresponders. New York: Guilford Press; 2003.

Bernstein DP, Cohen P, Skodol A, Bezirganian S, Brook JS. Childhood antecedents of adolescent personality disorders. American Journal of Psychiatry. 1996; 153:907–913. [PubMed: 8659613]

Blatt, SJ. Polarities of experience: Relatedness and self-definition in personality development, psychopathology, and the therapeutic process. American Psychological Association; 2008.

Blatt SJ, Lerner H. the psychological assessment of abject representation. Journal of Personality Assessment. 1983; 47:7–28. [PubMed: 6834235]

Bornovalova MA, Hicks BM, Iacono WG, McGue M. Stability, change, and heritability of borderline personality disorder traits from adolescence to adult- hood: A longitudinal twin study.

Development and Psychopathology. 2009; 21(Special Issue 4):1335–1353.10.1017/S0954579409990186 [PubMed: 19825271]

Bornstein RF. Behaviorally referenced experimentation and symptom validation: A paradigm for 21st century personality disorder research. Journal of Personality Disorders. 2003; 17:1–18. [PubMed: 12659543]

Bornstein RF. From symptom to process: How the *PDM* alters goals and strategies in Psychological Assessment. Journal of Personality Assessment. 2011; 93(2):142–150. [PubMed: 21347963]

Callahan KL, Price JL, Hilsenroth MJ. Psychological assessment of adult survivors of childhood sexual abuse within a naturalistic clinical sample. Journal of Personality Assessment. 2003; 80(2): 173–184. [PubMed: 12700020]

Caspi A, Roberts BW, Shiner RL. Personality development: Stability and change. Annual Review of Psychology. 2005; 56:453–484.

Cloninger, CR. The genetics and psychobiology of the seven-factor model of personality. In: Silk, KR., editor. Biology of personality disorders. Washington, DC: American Psychiatric Press; 1998.

Cohen P, Crawford T, Johnson J, Kasen S. The children in the community study of developmental course of personality disorder. Journal of Personality Disorders: Special Issue on Longitudinal Studies. 2005; 19(5):466–486.

Conway F, Oster M, McCarthy J. Exploring object relations in hospitalized children with caregiver loss. Journal of Infant, Child, and Adolescent Psychotherapy. 2010; 9(2-3):108–117.

Crawford T, Cohen P. Comorbid Axis I and Axis II disorders in early adolescence: Outcomes 20 years later. Archives of General Psychiatry. 2008; 65(6):641. [PubMed: 18519822]

Cross D, Westen D, Bradley B. Personality subtypes of adolescents who attempt suicide. Journal of Nervous and Mental Disease. 2011; 199:750–756. [PubMed: 21964268]

de Clercq B, van Leeuwen K, van den Noortgate W, de Bolle M, de Fruyt F. Childhood personality pathology: Dimensional stability and change. Development and Psychopathology. 2009; 21(Special Issue 3):853–869.10.1017/S0954579409000467 [PubMed: 19583887]

Defife JA, Drill R, Nakash O, Westen D. Agreement between clinician and patient ratings of adaptive functioning and developmental history. American Journal of Psychiatry. 2010; 167(2):1472–1478. [PubMed: 20634365]

DeFife JA, Goldberg M, Westen D. Dimensional assessment of self- and interpersonal functioning in adolescents: implications for DSM-5's general definition of personality disorder. Journal of Personality Disorders. in press.

DeMarree KG, Wheeler SC, Petty RE. Priming a new identity: Self-monitoring moderates the effects of nonself primes on self-judgments and behavior. Journal of Personality and Social Psychology. 2005; 89:657–671. [PubMed: 16351360]

Donegan NH, Sanislow CA, Blumberg HP, Fulbright RK, Lacadie C, Skudlarski P, et al. Wexler BE. Amygdala hyperactivity in borderline personality disorder: Implications for emotional dysregulation. Biological Psychiatry. 2003; 54:1284–1293. [PubMed: 14643096]

Dutra L, Campbell L, Westen D. Quantifying clinical judgment in the assessment of adolescent psychopathology: Reliability, validity, and factor structure of the Child Behavior Checklist for Clinician-Report. Journal of Clinical Psychology. 2004; 60:65–85. [PubMed: 14692010]

Ferguson CJ. A meta-analysis of normal and disordered personality across the life span. Journal of Personality and Social Psychology. 2010; 98(4):659–667. [PubMed: 20307136]

Ferguson, GA.; Takane, Y. Statistical analysis in psychology and education. Two correlation coefficients (n d). 6th1989. Retrieved March 27th, 2005, from http:/www.fon.hum.uva.nl/Service/Statistics/Two_Correlations.html

Field AP. Meta-analysis of correlation coefficients: A monte carlo comparison of fixed- and random-effects methods. Psychological Methods. 2001; 6:161–180. [PubMed: 11411440]

Gamache D, Diguer L, Laverdiere O, Rousseau JP. Development of an object relation-based typology of adolescent sex offenders. Bulletin of the Menninger Clinic. 2012; 76(4):329–364. [PubMed: 23244526]

Ganellen RJ. Assessing normal and abnormal personality functioning: Strengths and weaknesses of self-report, observer, and performance-based methods. Journal of Personality Assessment. 2007; 89:30–40. [PubMed: 17604532]

Greeno CG, Anderson CM, Shear MK, Mike G. Initial treatment engagement in a rural community mental health center. Psychiatric Services. 1999; 50(12):1634–1636. [PubMed: 10577887]

Gregory RJ, Mustata GT. Magical thinking in narratives of adolescent cutters. Journal of Adolescence. 2012; 25:1045–1051. [PubMed: 22464283]

Hall, GS. Adolescences: Its psychology and its relation to physiology, anthropology, sociology, sex, crime, religion, and education. Vol. I & II. Englewood Cliffs, NJ: Prentice-Hall; 1904.

Hall R. Global Assessment of Functioning: A modified scale. Psychosomatics. 1995; 36:267–275. [PubMed: 7638314]

Hilsenroth M, Ackerman S, Blagys M, Baumann B, Baity M, Smith S, Price J, Smith C, Heindselman T, Mount M, Holdwick D. Reliability and validity of the DSM-IV Axis V. American Journal of Psychiatry. 2000; 157:1858–1863. [PubMed: 11058486]

Hopwood CJ, Malone JC, Ansell EB, Sanislow CA, Grilo CM, McGlashan TH, et al. Morey LC. Personality assessment in the DSM-5: Empirical support for the rating severity, style, and traits. Journal of Personality Disorders. 2011; 25:305–320. [PubMed: 21699393]

Hull JG, Slone LB, Meteyer KB, Matthews AR. The nonconsciousness of self-consciousness. Journal of Personality and Social Psychology. 2002; 83:406–424. [PubMed: 12150237]

Huprich SK, Bornstein RF, Schmitt TA. Self-report methodology is insufficient for improving the assessment and classification of Axis II personality disorders. Journal of Personality Disorders. 2011; 25(5):557–570. [PubMed: 22023295]

Huprich SK, Greenberg RP. Advances in the assessment of object relations in the 1990s. Clinical Psychology Review. 2003; 23:665–698. [PubMed: 12971905]

Johnson JG, Cohen P, Skodol AE, Oldham JM, Kasen S, Brook JS. Personality disorders in adolescence and risk of major mental disorders and suicidality during adulthood. Archives of General Psychiatry. 1999; 56(9):805–811. [PubMed: 12884886]

Joint Commission on Accreditation of Healthcare Organizations. ORYX outcomes: The next evolution in accreditation: Performance measurement systems: Evaluation and selection. Oakbrook, IL: Author; 1997.

Jones SH, Thornicroft G, Coffey M, Dunn G. A brief mental health outcome scale:Reliability and validity of the Global Assessment of Functioning (GAF). British Journal of Psychiatry. 1995; 166:654–659. [PubMed: 7620753]

Kazdin, A. Child psychotherapy: Developing and identifying effective treatments. New York: Pergamon Press; 1988.

Kendall, P. Child and adolescent therapy: Cognitive-behavioral procedures. New York: Guilford Press; 1991.

Kernberg, O. Severe Personality Disorders: Psychotherapeutic Strategies. New Haven/London: Yale Univ. Press; 1984.

Kernberg O. Identity: Recent findings and clinical implications. Psychoanalytic Quarterly. 2006; 75:969–1004. [PubMed: 17094369]

Kihlstrom, JF.; Klein, SB. Self-knowledge and self-awareness. In: Snodgrass, JG.; Thompson, RL., editors. The self across psychology: Self-recognition, self-awareness, and the self concept. New York: New York Academy of Science; 1997. p. 5-17.

Klonsky ED, Oltmanns TF, Turkheimer E. Informant reports of personality disorder: Relation to self-reports, and future research directions. Clinical Psychology: Science and Practice. 2002; 9:300–311.

Kunda Z, Thagard P. Forming impression from stereotypes, traits, and behaviors: A parallel-constraint-satisfaction theory. Psychological Review. 1996; 103:284–308.

Linehan, M. Cognitive-behavioral treatment of borderline personality disorder. New York: Guilford; 1993.

Livesley WJ. A framework for integrating dimensional and categorical classifications of PD. Journal of Personality Disorders. 2007; 21:199–224. [PubMed: 17492921]

Livesley, WJ.; Jang, K.; Vernon, P. Genetic basis of personality structure. In: Millon, T.; Lerner, M., editors. Handbook of psychology: Personality and social psychology. Vol. 5. New York: John Wiley & Sons, Inc; 2003. p. 59-83.

Loevdahl H, Friis S. Routine Evaluation of Mental Health: Reliability information or worthless "guesstimates"? Acta Psychiatrica Scandinavica. 1996; 3:125–128. [PubMed: 8686482]

Lyons, J.; Howard, K.; O'Mahoney, M.; Lish, J. The measurement and management of outcomes in mental health. New York: Wiley; 1997.

McNamara, TP. Semantic priming: Perspectives from memory and word recognition. New York: Psychology Press/Taylor & Francis; 2005.

Markman KD, McMullen MN. A reflection and evaluation model of comparative thinking. Personality and Social Psychology Review. 2003; 7:244–267. [PubMed: 12788690]

Meyer GJ, Finn SE, Eyde LD, Kay GG, Moreland KL, Dies RR, Eisman EJ, Kubiszyn TW, Reed GM. Psychological testing and psychological assessment: A review of evidence and issues. American Psychologist. 2001; 56(2):128–165. [PubMed: 11279806]

Michels R, Siebel U, Freyberger HJ, Stieglitz RD, Schaub RT, Dilling H. The multiaxial system of the ICD-10: Evaluation of a preliminary draft in a multicentric field trial. Psychopathology. 1996; 29(6):347–356. [PubMed: 8994279]

Morey LC, Berghuis H, Bender DS, Verheul R, Krueger RF, Skodol AE. Toward a model for assessming level of personality functioning in DSM-%, Part II: Empirical articulation of a core dimension of personality pathology. Journal of Personality Assessment. 2011; 93:347–353. [PubMed: 22804673]

Murray HA. Thematic apperception test. 1943

Mussweiler T. Comparison processes in social judgment: Mechanisms and consequences. Psychological Review. 2003; 110:472–489. [PubMed: 12885111]

Nock MK. Self-injury. Annual Review of Clinical Psychology. 2010; 6:339–363.

Peters EJ, Hilsenroth MJ, Eudell-Simmons EM, Blagys MD, Handler L. Reliability and validity of the Social Cognition and Object Relations Scale in clinical use. Psychotherapy Research. 2006; 16(5): 617–626.

Pinsker J, Stein M, Hilsenroth M. The clinical utility of early memories as predictors of therapeutic alliance. Psychotherapy. 2007; 44:96–109. [PubMed: 22122172]

Podsakoff PM, MacKenzie SB, Lee JY, Podsakoff NP. Common method biases in behavioral research: A critical review of the literature and recommended remedies. Journal of Applied Psychology. 2003; 88(5):879–903. [PubMed: 14516251]

Porcerelli JH, Cogan R, Bambery M. The mental functioning axis of the psychodynamic diagnostic manual: An adolescent case study. Journal of Personality Assessment. 2011; 93(2):177–184. [PubMed: 21347966]

Porcerelli JH, Dauphin VB, Ablon JS, Leitman S, Bambery M. Psycho-analysis with avoidant personality disorder: A systematic case study. Psychotherapy: Theory/Research/Practice/Training. 2007; 44(1):1–13. [PubMed: 22122163]

Price JL, Hilsenroth MJ, Callahan KL, Petretic Jackson PA, Bonge D. A pilot study of psychodynamic psychotherapy for adult survivors of childhood sexual abuse. Clinical Psychology & Psychotherapy. 2004; 11(6):378–391.

Shevrin H, Dickman S. The psychological unconscious: A necessary assumption for all psychological theory. American Psychologist. 1980; 35:421–434.

Shiner RL. The development of personality disorders: Perspectives from normal personality development in childhood and adolescence. Development and Psychopathology. 2009; 21(Special Issue 3):715–734.10.1017/S0954579409000406 [PubMed: 19583881]

Shiner RL, Allen TA. Assessing personality disorders in adolescents: Seven guiding principles. Clinical Psychology: Science and Practice. 2013; 20:361–377.

Shiner, RL.; Tackett, JL. PDs in children and adolescents. In: Mash, EJ.; Barkley, RA., editors. Child Psychopathology. 3rd. New York, NY: Guilford Press; in press

Shrout PE, Fleiss JL. Interclass correlations: Uses in assessing rater reliability. Psychological Bulletin. 1979; 86(2):420–428. [PubMed: 18839484]

Skodol AE. PDs in *DSM-5*. Annual Review of Clinical Psychology. 2012; 8:317–334.

Skodol AE, Clark LA, Bender DS, Krueger RF, Morey LC, Verheul R, et al. Personality Disorders; Theory, Research, and Treatment. 2011; 2(1):4–22.

Soderberg P, Tungstrom S, Armelius BA. Reliability of the Global Assessment of Functioning ratings made by clinical psychiatric staff. Psychiatric Services. 2005; 56:434–438. [PubMed: 15812093]

Stein MB, Hilsenroth M, Pinsker-Aspen JH, Primavera L. Validity of DSM-IV Axis V Global Assessment of Relational Functioning Scale: A multimethod assessment. Journal of Nervous and Mental Disease. 2009; 197:50–55. [PubMed: 19155810]

Stein, M.; Hilsenroth, M.; Slavin-Mulford, J.; Pinsker, J. unpublished manuscript. Boston, MA: Massachusetts General Hospital and Harvard Medical School; 2011. Social Cognition and Object Relations Scale: Global Rating Method (SCORS-G). Retrieved January 8, 2013, from www.psychsystems.net/Manuals

Stein MB, Siefert CJ, Stewart RV, Hilsenroth MJ. Relationship between the Social Cognition and Object Relations Scale (SCORS) and attachment style in a clinical sample. Clinical Psychology & Psychotherapy. 2011; 18(6):512–523. [PubMed: 20806423]

Tackett JL, Balsis S, Oltmanns TF, Krueger RF. A unified perspective on personality pathology across the life span: Developmental considerations for the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders*. Development and Psychopathology. 2009; 21:687–713. [PubMed: 19583880]

Tracy K, Adler LA, Rotrosen J, Edson R, Lavori P. Interrater reliability issues in multicenter trials, part 1: Theoretical concepts and operational procedures used in Department of Veterans Affairs Cooperative Study #394. Psychopharmacology Bulletin. 1997; 33(1):53–57. [PubMed: 9133751]

Tuma, JM. Traditional therapies with children. In: Ollendick, T.; Hersen, M., editors. Handbook of child psychopathology. New York: Plenum Press; 1983. p. 419-437.

Turner MA, Boden JM, Mulder RT. Predictors of hospitalization two years after treatment for first-episode psychosis. Psychiatric Services. 2013; 64(12):1230–1235. [PubMed: 24036506]

Wagner AW, Linehan MM. Facial expression recognition ability among women with borderline personality disorder: Implications for emotion regulation? Journal of Personality Disorders. 1999; 13:329–344. [PubMed: 10633314]

Westen D. Clinical assessment of object relations using the TAT. Journal of Personality Assessment. 1991; 56:56–74. [PubMed: 16370902]

Westen, D. Unpublished manuscript. Department of Psychiatry, The Cambridge Hospital and Harvard Medical School; Cambridge, MA: 1995. Social Cognition and Object Relations Scale: Q-Sort for Projective Stories (SCORS-Q).

Westen D, Betan E, DeFife JA. Identity disturbance in adolescence: Associations with borderline personality disorder. Development and Psychopathology. 2011; 23:305–313. [PubMed: 21262056]

Westen D, Gabbard GO. Developments in cognitive neuroscience: I. Conflict, compromise, and connectionism. Journal of the American Psychoanalytic Association. 2002; 50:53–98. [PubMed: 12018875]

Westen D, Klepser J, Ruffins S, Silverman M, Lifton N, Boekamp J. Object relations in childhood and adolescence: The development of working representations. Journal of Consulting and Clinical Psychology. 1991; 59:400–409. [PubMed: 2071725]

Westen D, Lohr N, Silk K, Gold L, Kerber K. Object relations and social cognition in borderlines, major depressives, and normals: A TAT analysis. Psychological Assessment: A Journal of Consulting and Clinical Psychology. 1990; 2:355–364.

Westen D, Ludolph P, Lerner H, Ruffins S, Wiss FC. Object relations in borderline adolescents. Journal of the American Academy of Child and Adolescent Psychiatry. 1990; 29:338–348. [PubMed: 2347828]

Westen D, Ludolph P, Misle B, Ruffins S, Block MJ. Physical and sexual abuse in adolescent girls with borderline personality disorder. American Journal of Orthopsychiatry. 1990; 60:55–66. [PubMed: 2305845]

Westen D, Muderrisoglu S, Fowler C, Shedler J, Koren D. Affect regulation and affective experience: Individual differences, group differences, and measurement using a Q-sort procedure. Journal of Consulting and Clinical Psychology. 1997; 65:429–439. [PubMed: 9170766]

Westen D, Shedler J, Durrett C, Glass S, Martens A. Personality diagnosis in adolescence: DSM-IV Axis II diagnoses and an empirically derived alternative. American Journal of Psychiatry. 2003; 160:952–966. [PubMed: 12727701]

Wheeler, SC.; DeMarree, KG.; Petty, RE. The roles of the self in priming-to-behavior effects. In: Tesser, A.; Wood, JV.; Stapel, DA., editors. On building, defending and regulating the self: A psychological perspective. New York: Psychology Press; 2005. p. 245-271.

Whipple R, Fowler JC. Affect, relationship schemas, and social cognition: Self-injuring Borderline Personality Disorder inpatients. Psychoanalytic Psychology. 2011; 28(2):183–185.

Winters NC, Collett BR, Myers KM. Ten year review of rating scales, VII: Scales assessing functional impairment. Journal of the American Academy of Child and Adolescent Psychiatry. 2005; 44(4): 309–338. [PubMed: 15782079]

Young, J. Cognitive therapy for personality disorders: A schema-focused approach. Sarasota, Fl: Professional Resources Exchange; 1990.

Zeeck A, Hartmann A, Orlinsky DE. Internalization of the therapeutic process: Differences between borderline and neurotic patients. Journal of Personality Disorders. 2006; 20:22–41. [PubMed: 16563077]

Table 1
Means and standard deviations for study measures

	N	Mean	SD	Range
SCORS-G composite*	66	3.28	.59	2.14-5.07
Global Rating of Engagement in Group Therapy	65	2.71	1.11	1-5
Global Rating of Engagement in Individual Therapy	57	3.28	1.19	0-5
GAF*	67	41.51	6.18	27.5-57.5
Chronic Level of Personality Functioning	67	2.7	.65	1-4
Quality of Peer Relationships	67	2.70	.78	1-4
School Functioning	66	2.48	.90	1-5

Note: SCORS-G = Social Cognition and Object Relations Scale- Global Version; GAF = Global Assessment of Functioning

* was calculated averaging together the individual and group therapist ratings.

Table 2

Correlational Matrix between SCORS-G items and Study Variables

SCORS-G Items	ICC[1]	Engagement Group Therapy	Engagement Individual Therapy	Admission GAF	Personality Functioning	Peer Relationships	School Functioning	NSSI History	Eating Disorder History
COM	.65	.40**	.42**	.55**	.27*	.07	.42**	-.17	-.19
AFF	.38	.28*	.16	.37**	.31*	.29*	.30*	-.09	.04
EIR	.56	.28*	.36**	.56**	.46**	.58**	.46**	-.26*	-.20
EIM	.65	.23	.51**	.45**	.64**	.28*	.55**	-.38**	-.33**
CS	.63	.34**	.57**	.57**	.34**	.32*	.45**	-.26*	-.24*
AGG	.66	.22	.25	.29*	.51**	.43**	.45**	-.13	-.07
SE	.17	.41**	.14	.36**	.20	-.11	.19	-.14	-.07
ICS	.35	.06	.17	.24	.19	.17	.21	-.07	-.05
TOT	.64	.41**	.48**	.57**	.49**	.30*	.54**	-.29*	-.31**

* $p < .05$;

** $p < .01$

Note:

[1] = ICC (1,2); NSSI = Non-suicidal self-injury; COM = Complexity of Representations of Self and Others; AFF = Affective Quality of Representations; EIR = Emotional Investments in Relationships; EIM = Emotional Investment in Values and Moral Standards; CS = Understanding Social Causality; AGG = Experience and Management of Aggressive Impulses; SE = Self-Esteem; ICS = Identity and Coherence of Self; TOT= total composite score of all items

Table 3

Hierarchical Regression of the Incremental Validity of the SCORS-G Composite Score in Predicting Composite Therapeutic Engagement Rating with GAF Scores Entered First (Model 1) and with GAF Scores Entered After SCORS-G Ratings (Model 2).

Model	Block	Variable Entered	*B*	*SE*	*β*	R^2	*F* for change in R^2
1	1	GAF	.06	.02	.37**	.14	
	2	GAF	.03	.02	.17	.24	6.95*
		SCORS-G	.60	.23	.37*		
2	1	SCORS-G	.76	.19	.46	.22	
	2	SCORS-G	.60	.23	.37	.24	1.51
		GAF	.03	.02	.17		

* $p<.05$,

** $p<.01$;

$N = 57$

Table 4

Hierarchical Regression of the Incremental Validity of the SCORS-G Composite Score in Predicting Overall Functioning Rating with GAF Scores Entered First (Model 1) and with GAF Scores Entered After SCORS-G Ratings (Model 2).

Model	Block	Variable Entered	*B*	*SE*	*β*	R^2	*F* for change in R^2
1	1	GAF	.01	.04	.35	.12	7.87
	2	GAF	.01	.04	.03	.34	18.28**
		SCORS-G	1.66	.39	.57**		
2	1	SCORS-G	1.72	.32	.58**	.34	28.990
	2	SCORS-G	1.66	.39	.57**	.34	.07
		GAF	.03	.02	.17		

*$p<.05$,

**$p<.01$;

$N = 57$

Table 5
Correlational Matrix Between Individual and Group Therapists Ratings

	Group Therapists' Rated		
Individual Therapists' Rated	**SCORS-G**	**Engagement in Group Therapy**	**Admission GAF**
SCORS-G	**.50****	**.30***	**.42****
Engagement in Individual Therapy	**.30***	**.43****	**.26***
Admission GAF	**.31***	.12	**.40****
Chronic Level of Personality Functioning	.24	-.15	**.33****
Peer Relationships	.15	-.07	.14
School Functioning	**.36****	.21	.23

* $p < .05$;

** $p < .01$

Table 6

Focused Effect Size Comparisons Between SCORS-G correlations with Criterion Variables from This Study and Dataset from Defife et al., 2013

	This Study (n = 60)	Defife et al. (2013) Dataset (n = 254)	Focused Effect Size Comparison	
	r	*r*	*z*	*p*
Chronic Level of Personality Functioning	.49	.53	.37	.71
Peer Relationships	.30	.51	1.73	.08
School Functioning	.54	.45	.81	.42