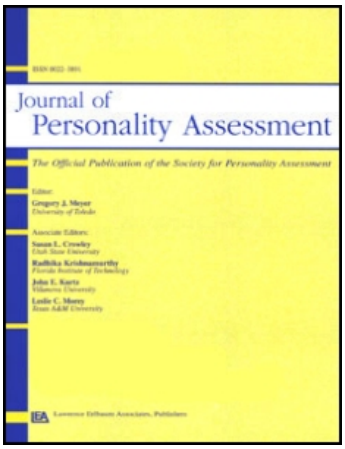

Journal of Personality Assessment
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t775653663

An Examination of the Construct Validity of the Rorschach Mutuality of Autonomy (MOA) Scale

George Bombel [a]; Joni L. Mihura [a]; Gregory J. Meyer [a]
[a] Department of Psychology, University of Toledo,

Online Publication Date: 01 May 2009

To cite this Article Bombel, George, Mihura, Joni L. and Meyer, Gregory J.(2009)'An Examination of the Construct Validity of the Rorschach Mutuality of Autonomy (MOA) Scale',Journal of Personality Assessment,91:3,227 — 237
To link to this Article: DOI: 10.1080/00223890902794267
URL: http://dx.doi.org/10.1080/00223890902794267

Journal of Personality Assessment, *91*(3), 227–237, 2009

ISSN: 0022-3891 print / 1532-7752 online
DOI: 10.1080/00223890902794267

Routledge
Taylor & Francis Group

ARTICLES

An Examination of the Construct Validity of the Rorschach Mutuality of Autonomy (MOA) Scale

GEORGE BOMBEL, JONI L. MIHURA, AND GREGORY J. MEYER

Department of Psychology, University of Toledo

Using 100 clinical cases, we examined the construct validity of the Mutuality of Autonomy (MOA) Scale (Urist, 1977) using Westen and Rosenthal's (2003) $r_{contrast\ -\ construct\ validity\ (CV)}$ procedure for quantifying a pattern of convergent-discriminant relationships between a target measure and a set of criterion variables. Our 15 criterion variables included the Comprehensive System (CS; Exner, 2003) variables, a CS-based measure of ego strength (Resnick, 1994), and 3 subscales from the Social Cognition and Object Relations Scale (Westen, Lohr, Silk, Kerber, & Goodrich, 1990). We generated the $r_{contrast\ -\ CV}$ coefficients to test 2 competing hypotheses: that the MOA Scale primarily measures object relations (OR) quality or that it primarily measures psychopathology. Results suggest that the MOA Scale is an equally potent measure of OR and psychopathology regardless of the MOA Scale index used.

Object relations (OR) constructs are important subjects of study in personality assessment research. Literature reviews dedicated to the subject have been published routinely (Fishler, Sperling, & Carr, 1990; Frank, 1995; Huprich & Greenberg, 2003; T. E. Smith, 1993; Stricker & Healey, 1990). One of the most popular OR assessment instruments is the Rorschach Inkblot Method (Rorschach, 1921/1942), for which a number of OR scoring systems have been developed. Some of these systems are rarely used today (Coonerty, 1986; Kwawer, 1979; Mayman, 1967; Pruitt & Spilka, 1964). Others continue to make appearances in the literature (Blatt, Brenneis, Schimek, & Glick, 1976; Burke, Friedman, & Gorlitz, 1988; Cooper, Perry, Hoke, & Richman, 1985). The scoring system by Urist (1977; Mutuality of Autonomy [MOA] Scale) is one of the most well-known among Rorschach OR assessment methods (Huprich & Greenberg, 2003; Stricker & Healey, 1990). The MOA Scale is an especially attractive measure because it is relatively simple to use; for example, scorers who are relatively unfamiliar with OR theory can score it reliably (Holaday & Sparks, 2001). Furthermore, unlike other measures, Rorschach responses with inanimate, nonhuman-like figures are scorable with the MOA Scale.

Urist (1977) developed the MOA Scale as a measure of object relational developmental maturity or the degree to which individuation separation has been attained. The seven points, or levels, of this scale represent specific developmental levels along a dimension from empathic, reciprocal relatedness to destructive envelopment and symbiotic fusion. Each level, more specifically, corresponds to "developmentally significant gradations in the individual's capacity to experience self and other as mutually autonomous within relationships" (Urist, 1977, p. 4). Synopses of the scoring guidelines, for the healthiest to most primitive levels, are the following:

1. Figures are described as engaging in an activity or relationship that conveys a sense of mutuality, with a reciprocal acknowledgment of their respective individuality (e.g., "Two griffons building a nest together.").
2. Figures are described as engaging in a relationship or parallel activity without emphasis on reciprocity or mutuality, although the description does not necessarily need to compromise mutuality (e.g., "Two bears climbing up different sides of the same mountain.").
3. Figures are described as engaging in an activity in which they require each other, or one requires the other, that is, an external source of support or direction is required (e.g., "A clown hanging from a cliff.").
4. The described relationship conveys a sense that the definition or stability of one object necessarily requires the other because it is merely an extension or reflection of the other (e.g., "A woman scrutinizing her reflection in the mirror.").
5. The described relationship themes are characterized by severe imbalances of mutuality in which figures are powerless and helpless versus omnipotent and controlling. Themes of influence and controlling are present (e.g., "A sorcerer casting a spell on someone.").
6. The figure(s) are described as engaging in activity that is clearly destructive or parasitic, and that compromises the autonomy or integrity of the victim (e.g., "A giant squashing a dwarf.").
7. The described relationship theme is envelopment and engulfment in which one figure has devoured or consumed another, either explicitly or implicitly (e.g., "It's an X-ray of a creature that's eaten another creature.").

Received July 19, 2007; Revised September 21, 2008.

Editor's Note: Radhika Krishnamurthy served as the Editor with full decision authority over this manuscript.

Address correspondence to George Bombel, UTHSCSA, Department of Psychiatry, 7703 Floyd Curl Drive, MC 7792, San Antonio, TX 78229–3900; Email: gbombel@gmail.com

Importantly, Urist (Urist & Shill, 1982) made a distinction between the phenomena he was attempting to measure—OR development—and "a more general Rorschach health-sickness factor" (p. 453). Indeed, in Urist and Shill's second MOA Scale publication, they took explicit methodological steps to ensure that the latter would not influence raters' efforts to record the former in participants' Rorschach responses. Theoretically, this distinction is a valid one. Object relational development is the maturational process of separation individuation that can lead to healthy outcomes (Masterson, 1976). "Health-sickness," or psychopathology, does not refer to a developmental concept or personality organization but rather to the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., text revision [*DSM–IV–TR*]; American Psychological Association, 2000) Axis I psychopathology (e.g., dysphoria, phobic avoidance, thought disorder). The constructs, therefore, are not synonymous. Still, when developmental outcomes are unhealthy (e.g., borderline personality organization), vulnerability to developing and expressing psychopathology presumably increases (Berg, Packer, & Nunno, 1993). In these cases, object representations of poor quality are likely to co-occur with symptomatology.

The MOA Scale has respectable psychometric qualities. However, Urist's (1977) initial interrater reliability values were not impressive: Raters achieved exact agreement at the response level only 52% of the time and were within 1 point of each other 86% of the time. Holaday and Sparks (2001) calculated the unweighted response-level exact agreement average in the literature to be 74.3%. After expanding the MOA Scale scoring guidelines, Holaday and Sparks obtained an exact agreement rate of 80% at the response level using raters naïve to OR theory. Results from a recent meta-analysis of MOA Scale interrater scoring reliability (Bombel, 2006) found weighted kappa and percent agreement coefficients of .83 and .81, respectively, when these statistics were computed across individual responses, whereas intraclass correlations (ICCs) and Pearson's *r* s were .94 and .91, respectively, when computed at the protocol level, that is, when raters were compared across aggregated protocol scores. These data indicate that the MOA Scale can be reliably scored.

The MOA Scale has also demonstrated validity as a measure of OR quality. In Urist's (1977) initial explication of the measure, MOA Scale scores correlated significantly with (a) inpatients' autobiographical material scored for object relational themes ($r = .63$, $p < .001$) and (b) staff ratings of the quality of the inpatients' relationships with hospital staff and other inpatients ($r = .43$, $p < .001$). Among an inpatient sample, Blatt, Tuber, and Auerbach (1990) found significant correlations ($r = .26$, $p < .001$) between MOA Scale scores and the Concept of the Object Scale (Blatt et al., 1976), which assigns OR scores to thematic and structural aspects of Rorschach responses. Ryan, Avery, and Grolnick (1985), in an investigation of the convergent and discriminant validity of the MOA Scale, found that healthier (i.e., lower) MOA Scale responses were significantly related to social adjustment ($r = -.24$, $p < .10$), self-esteem ($r = -.26$, $p < .05$), and interpersonal cooperation ($r = -.33$, $p < .05$) but not to teacher ratings of achievement and intelligence. Finally, a recent meta-analysis (Fowler, Addelson, & Clemence, 2006) of the validity of the MOA Scale yielded small to medium effect sizes (*r*s = .20–.34) across several indexes when criteria were correlates of OR phenomena, such as observer-rated behaviors, and/or self-reports of violence, marital problems, and relationship functioning.

At the same time, some data suggest that the MOA Scale does not cleanly measure what it purports to measure. Tuber (1992) noted that the scale does not correlate with chronological age, which means that its scoring levels do not represent points on an OR developmental timeline but simply object representations of varying qualities. Furthermore, the scale may also be a potent measure of psychopathology. Harder, Greenwald, Wechsler, and Ritzler (1984) found that higher (more pathological) MOA mean scores were significantly associated with diagnosis severity trichotomized into the categories of nonpsychotic, affective psychotic, and schizophrenic spectrum disorders ($r = .43$, $p < .001$).

In a thorough study with long-term psychiatric inpatients, Blatt et al. (1990) investigated the relationships between several MOA Scale scores (e.g., the mean, single healthiest score, and the single most pathological score) and (a) estimates of social behavior and interpersonal relations derived from clinical records, (b) neurotic and psychotic symptomatology as recorded in clinical records, and (c) Rorschach measures of reality testing (F + %) and thought disorder (e.g., severity of boundary disturbance) developed earlier by Blatt and Ritzler (1974). Surprisingly, although most adaptive scores related to a measure of interpersonal communication, MOA Scale mean scores correlated significantly with symptom severity ($r = .35$, $p < .0001$), reality testing ($r = -.41$, $p < .0001$), and thought disorder ($r = .49$, $p < .0001$) but not with ratings of social behavior and interpersonal relations on the ward. MOA Scale indexes of pathological object representations (i.e., the sum of responses scored 5, 6, or 7) evidenced a similar pattern of significant and nonsignificant results. Blatt et al. (1990) concluded that the MOA Scale appears to measure psychopathology primarily and interpersonal relationship quality secondarily.

Berg et al. (1993) investigated the relationship between thought disorder and object representation in psychiatric outpatients diagnosed with borderline personality disorder, narcissistic personality disorder, and schizophrenia. A strong relationship was found between MOA Scale pathology scores (sum of Levels 5, 6, and 7) and thought disorder ($r = .58$, $p < .001$) as measured by the Comprehensive System (CS; Exner, 1986) weighted sum of the six cognitive special scores (WSum6) variable. Further, whereas the frequencies of MOA Scale Level 4 and 5 scores differentiated the patients with personality disorders from those with schizophrenia, the pathology composite did not successfully discriminate among the three patient groups. These data suggest that the MOA Scale is not making finer distinctions between poor object representations and thought disturbances when both are present.

The results from these studies highlight concerns that have been raised in literature reviews about the construct validity of the MOA Scale (Stricker & Healey, 1990; Tuber, 1992). It appears to be sensitive to relationship quality and social functioning; however, it may be no less sensitive to severity of Axis I psychopathology. If this is the case, MOA Scale scores are confounded and difficult to interpret. Given the popularity of the scale, a construct validity investigation exploring this issue is warranted.

Cronbach and Meehl (1955) noted that one could estimate the construct validity of a measure by providing evidence to show that it interrelates with measures of other constructs in ways that

would be expected if the underlying theory is true. Recently, Westen and Rosenthal (2003) developed a procedure to quantify the pattern of relationships in this conceptual framework. Based on analysis of variance contrast analysis, it involves generating a set of predicted relationships (i.e., a pattern of contrast weights) between the target measure and the specified criterion variables and then calculating the degree to which this prediction matches the actual relationship pattern. In the true spirit of contrast analysis, the predicted relationship pattern between the measure and the criterion variables constitutes the hypothesis to be tested. The resulting effect size, $r_{contrast\text{ - }construct\ validity\ (CV)}$, is a Pearson's product–moment correlation coefficient that is readily interpretable and readily aggregated in a future meta-analysis (Westen & Rosenthal, 2003).

To investigate the construct validity of the MOA Scale, Westen and Rosenthal's (2003) procedure was used in this study to test competing hypotheses about what the scale actually measures. That is, Westen and Rosenthal's procedure was conducted in two different circumstances resulting in two separate $r_{contrast\text{ - }CV}$ values—the first assuming that the MOA Scale is a measure of OR and the second assuming the MOA Scale is a measure of psychopathology. Therefore, first, we created a set of predicted correlations between the MOA Scale and our criterion variables assuming that the scale is primarily a measure of the quality of OR. For example, the MOA Scale should correlate more strongly with other measures of OR and less strongly with measures of thought disorder and Axis I symptom severity. We then used the $r_{contrast\text{ - }CV}$ procedure to generate an effect size representing the degree to which this MOA Scale as OR measure prediction set matched the actual pattern of relationships between the MOA Scale and the 14 criterion variables chosen for this study. Next, we created a set of predicted correlations between the MOA Scale and the criterion variables assuming that the scale is primarily a measure of Axis I symptom severity. This time, the $r_{contrast\text{ - }CV}$ procedure generated an effect size representing the degree to which this MOA Scale as pathology measure prediction set matched the actual pattern of relationships between the MOA Scale and the criterion variables. The competing prediction sets yield the following expectations: (a) If the MOA Scale is primarily an OR measure, the $r_{contrast\text{ - }CV}$ effect size for the MOA Scale as OR measure prediction set should be larger than the $r_{contrast\text{ - }CV}$ effect size for the MOA Scale as pathology measure prediction set; (b) contrariwise, if the MOA Scale is primarily a measure of psychopathology severity, the $r_{contrast\text{ - }CV}$ effect size for the MOA Scale as pathology measure prediction set should be larger than the $r_{contrast\text{ - }CV}$ effect size for the MOA Scale as OR measure prediction set.

We tested this set of competing hypotheses four times in our study, once for each of our MOA Scale score indexes. The first three are (a) MOA Scale mean score (MOAx), which is reported frequently in the literature (Stricker & Healey, 1990); (b) the single best or most adaptive score (MOAb); and (c) the single worst or most pathological score (MOAp). These latter two scores have also been reported very frequently in the literature (e.g., Ackerman, Hilsenroth, Clemence, Weatherill, & Fowler, 2000; Blais, Hilsenroth, Fowler, & Conboy, 1999; Cook, Blatt, & Ford, 1995; Strauss & Ryan, 1987; Tuber, 1983). The final measure was a refined variable that makes use of the full spectrum of MOA scores defined here as (d) the relative Health index (MOAHI) and computed as the number of healthy Level 1 scores minus the most disturbed Level 5, 6, and 7 scores divided by the total number of scorable MOA Scale responses. This score is a ratio of developmentally healthy to primitive scores relative to the number of MOA responses produced. Thus, it provides a more sophisticated summary of MOA Scale information than the other indexes.

Method

Participants

The participants in this study were drawn from a large archive ($N = 440$) of individuals who underwent psychological assessment through a testing service associated with a large Midwestern university hospital. To be considered, participants' case files needed to include a Rorschach protocol with between 14 and 50 responses (R). Limits were placed on R to help ensure results would not be negatively affected by unreliably brief or excessively lengthy protocols (Dean, Viglione, Perry, & Meyer, 2007; Exner, 2003; Meyer, 1992). Also, case files needed to include Thematic Apperception Test (TAT; H. A. Murray, 1943) stories for cards 1, 2, 3BM, 4, 10, and 13MF. We chose the first 100 case files meeting these criteria, starting from the most recently created, for analyses here. The average number of Rorschach responses in the final sample was 22.71 ($SD = 8.15$). This group was comprised of psychiatric inpatients ($N = 30$), psychiatric outpatients ($N = 36$), general medical inpatients and outpatients evaluated for psychiatric reasons ($N = 33$), and 1 incarcerated individual evaluated for psychiatric reasons. More males ($N = 58$) than females ($N = 42$) were represented, and age of the participants ranged from 15 to 71 years ($M = 34.5$, $SD = 12.99$). Years of education ranged from 9 to 21 ($M = 13.74$, $SD = 2.55$). At the time of testing, 28 of the individuals were married, 10 were divorced, 3 were widowed, and 59 had never been married.

Measures

The criterion variables used in this study were as follows:

1. The Complexity of Human Representations (CHR) scale from the Social Cognition and Object Relations Scale (SCORS; Westen, Lohr, Silk, Kerber, & Goodrich, 1990).
2. The Affect Tone (AT) scale from the SCORS.
3. The Emotional Investment in Human Relationships (EIR) scale from the SCORS-Global Method scale (SCORS-G; Hilsenroth et al., 2004).

and the following variables from the CS (Exner, 2003):

4. Good Human Representation (GHR).
5. Poor Human Representation (PHR).
6. Human Representational Variable (HRV).
7. X – %.
8. WDA%.
9. WSum6.
10. Perceptual Thinking Index (PTI).
11. Suicide Constellation (S-CON).
12. Depression Index (DEPI).
13. Coping Deficit Index (CDI).
14. A composite scale derived from CS scores called the Conceptual Ego Strength Index (CESI; Resnick, 1994; Resnick & Meyer, 1995).

These variables were included because we hypothesized that they would relate to the target constructs (OR or psychopathology) in varying degrees, and the $r_{contrast - CV}$ procedure is most sensitive to levels of agreement between the predicted and observed correlations when the correlations include a wide range of values (Westen & Rosenthal, 2003). Further, we made an effort to include criterion variables that assess aspects of psychopathology associated with both the more neurotic (e.g., DEPI) and more psychotic (e.g., WSum6) segments of the psychoanalytic personality organization spectrum (Kernberg, 1986; McWilliams, 1994).

SCORS. The SCORS, and its briefer version, the SCORS–G, are eight scale scoring systems designed to broadly assess affective processes, social information processing, and internalized representations of self and others. Interrater reliability of the SCORS has been well documented when used to rate TAT stories (Ackerman et al., 2000; Fowler et al., 2004), early memories (Fowler, Hilsenroth, & Handler, 1995), and narrative statements by patients during psychotherapy sessions (Peters, Hilsenroth, Eudell-Simmons, Blagys, & Handler, 2006). Further, the SCORS has demonstrated convergent and construct validity by relating significantly to other OR measures such as the MOA Scale (Ackerman, Hilsenroth, Clemence, Weatherill, & Fowler, 2001) and the Concept of the Object Scale (Hibbard, Hilsenroth, Hibbard, & Nash, 1995).

In this study, we used the CHR and AT scales from the SCORS and the EIR scale from the SCORS–G as criterion variables because they were expected to converge with the MOA Scale to differing degrees. The CHR scale assesses the capacity for seeing self and others as separate individuals with enduring dispositions, complex motives, and unique subjective experiences. The AT scale measures the emotional tone of what an individual expects from others and the world, ranging from pain and malevolence to positive, enriching, and happy experiences. The EIR scale assesses the quality of relationship representations, ranging from profound selfishness to mutuality, respect, and interdependence. We used Westen et al.'s (1990) 5-point CHR and AT scales here because they provide ample guidelines for rater training. The EIR scale from the SCORS–G was chosen instead of the Emotional Investment in Relationships and Moral Standards scale from the SCORS because it focuses more on relational phenomena. Typically, final ratings for each scale are generated by averaging scale scores across individual TAT stories (or early memories, etc.).

Rorschach CS. The Rorschach criterion variables used in this study are in the CS (Exner, 2003) or derived from its scores. The original Human Experience Variable (HEV), which included the Good Human Experience (GHE) and Poor Human Experience (PHE) variables, was recently modified to create the psychometrically superior trio of variables mentioned earlier: the HRV, GHR, and PHR (Viglione, Perry, Jansak, Meyer, & Exner, 2003). Viglione et al. (2003) described the new variables as follows:

> GHR responses are perceptions or representations of positive schema of self, other, and relationships manifested in accurate, realistic, logical, intact, human responses and benign or cooperative interactions. PHR are negative and problematic perceptions or representations as manifested in distorted, unrealistic, damaged, confused, illogical, aggressive, or malevolent representation or perceptions. Along these broad and heterogeneous dimensions, the HRV summarizes the overall quality of human interpersonal perceptions and representations, that is, the implicit understanding of people and relationships. (p. 71)

Research using the newer HRV measures has been limited to just a few studies. However, Viglione et al. (2003) found substantial correlations between the HEV measures and their HRV counterparts (all rs $\geq .87$), and the HEV has demonstrated construct validity. For example, Burns and Viglione (1996) found that HEV distinguished between high and low interpersonal relatedness groups as identified with the Bell Object Relations Inventory (BORI; Bell, 1991), a version of the BORI modified to be a spouse-report instrument, and the Emotional Maturity Rating Form (Bessell, 1984).

The X – % variable represents the proportion of Rorschach responses in which the individual disregarded the contours of the inkblots when formulating percepts. Exner's (1993) reference group of individuals with schizophrenia had higher X – % scores than the nonpatient normative group. Similarly, Dao and Prevatt (2006) found that depressed individuals without psychosis scored significantly lower (healthier) on the X – % variable than did individuals with a schizophrenia-spectrum disorder. WDA% represents the extent to which a respondent is using common blot contours in conventional ways. In the aforementioned study by Dao and Prevatt, the depressed group scored significantly higher (more conventional) than the schizophrenia-spectrum group.

The WSum6 indicates the presence of thinking disturbance or a thought disorder (Exner, 2003). In Dao and Prevatt's (2006) sample, the schizophrenia-spectrum group exhibited significantly higher WSum6 scores than did the depressed group. Further, it has differentiated children and adolescents with disturbed thinking and behavior from those without such problems (S. R. Smith, Baity, Knowles, & Hilsenroth, 2001).

The S-CON is an actuarial index that helps identify individuals who may be at risk for suicide (Exner, 2003). Fowler, Piers, Hilsenroth, Holdwick, and Padawer (2001) found that S-CON scores ≥ 7 predicted near-lethal suicide attempts among a mixed group of psychiatric patients and successfully discriminated between such attempts and parasuicidal behavior.

Exner (2003) suggested that the DEPI is sensitive to some aspects of depression, although not necessarily a *DSM* major depressive illness. Still, in a review and meta-analysis by Jørgenson, Andersen, and Dam (2000), the DEPI was more able to identify "true positives" when the diagnoses were nonpsychotic and unipolar depression than when target diagnoses were psychotic and bipolar depression and when depression was associated with borderline personality disorder. Jørgenson et al. noted that the DEPI may be more sensitive to unambiguous forms of depression; however, its overall effect size as a diagnostic measure of depression appears to be small.

The CDI is comprised of a heterogeneous mix of variables related to coping resources, capacity for directing responses to environmental challenges, and interpersonal difficulties (Exner, 2003). Higher scorers tend to be socially inept or helpless and have impoverished and unfulfilling interpersonal relationships. Exner (2003) described an unpublished study wherein CDI scores differentiated outpatients who reported interpersonal complaints from those that did not.

Resnick (1994) developed the CESI to measure the ego psychological construct of ego strength. It is based on Kleiger's

(1992) notion that the EA variable from the CS can be interpreted as an indicator of ego-mediated coping resources, including the capacity to endure stress, anxiety, and other negative affect states. The CESI is computed from CS scores for Form dominance, Form quality, cognitive Special Scores, and primitive Content. Meyer and Resnick (1996; Resnick & Meyer, 1995) have found that even though the CESI has a strong correlation ($r = .72$) with the Ego Impairment Index (EII; Perry & Viglione, 1991), it slightly outperformed the EII by correlating more strongly with clinician-assigned *DSM* diagnostic codes along a theoretical continuum of ego impairment ($rs = .40$ vs. .35). It also correlated more strongly with the criterion than the Ego-Strength Scale (Barron, 1953) and Goldberg Index (Goldberg, 1965) from the Minnesota Multiphasic Personality Inventory–2 (Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989; rs of –.25 and .17, respectively).

Importantly, there are no specific scoring confounds between the Rorschach criterion variables used here and the MOA Scale. All of the Rorschach criterion variables are derived primarily from structural, or quantitative, aspects of responses across protocols. For example, although GHR, PHR, the CDI, and the CESI are sensitive in varying degrees to content themes of aggression and/or cooperation (like the MOA Scale), they are calculated largely from specific structural summary elements (e.g., FQ, Special Scores, noninterpersonal theme Content) in the protocol (Exner, 2003). In contrast, MOA Scale scores are assigned based only on molar, interpersonally oriented thematic elements; their values do not represent formal aspects of percepts but rather OR developmental levels (Urist, 1977).

TABLE 1.—Interrater reliability coefficients and descriptive statistics for the MOA Scale indexes and criterion variables.

Measure	ICC[a]	Mean	SD	Minimum	Maximum
MOAx	.91	2.87	0.78	1.00	5.00
MOAb	.75	1.57	0.67	1.00	4.00
MOAp	.88	4.62	1.59	1.00	7.00
MOAHI	.78	−0.05	0.25	−0.71	1.00
CHR	.59	2.94	0.60	2.00	4.67
AT	.71	2.92	0.49	1.50	4.08
EIR	.69	3.38	0.60	1.83	5.67
GHR	.92	3.33	1.96	0.00	9.00
PHR	.96	4.55	4.04	0.00	18.00
HRV	.90	−1.22	4.34	−15.00	7.00
X - %	.87	0.21	0.12	0.00	0.56
WDA%	.91	0.81	0.12	0.50	1.00
WSum6	.81	23.13	25.84	0.00	162.00
PTI	.78	1.30	1.46	0.00	5.00
S-CON	.76	5.82	1.81	1.00	10.00
DEPI	.83	4.53	1.37	1.00	7.00
CDI	.89	2.96	1.31	0.00	5.00
CESI	.86	0.53	0.21	0.14	1.32

Note. Descriptive data based on $N = 100$. MOA = Mutuality of Autonomy Scale; ICC = intraclass correlation: 2-way random effects model absolute agreement ICC (A,1; McGraw & Wong, 1996) computed at the summary-score level; MOAx = MOA Scale mean; MOAb = MOA Scale best score index; MOAp = MOA Scale Pathology index; MOAHI = MOA Scale Health index; SCORS = Social Cognition and Object Relations Scale; CHR = Complexity of Human Representation (CHR) SCORS scale; AT = Affect Tone (AT) SCORS scale; EIR = Emotional Investment in Human Relationships (EIR) SCORS scale; PTI = Perceptual Thinking index; S-CON = Suicide Constellation; DEPI = Depression index; CDI = Coping Deficit index; CESI = Conceptual Ego Strength index.
[a] CHR, AT, and EIR based on $N = 20$; all other variables based on $N = 23$.

Procedures

Scoring and reliability. All Rorschach tests in the original database were administered and scored by G. J. Meyer or one of several clinical psychology doctoral students in training with him who had received course work and supervised training in the CS before data collection began. Protocols were scored with the CS at the time of testing. To determine interrater reliability, a batch of randomly selected protocols ($N = 23$) was independently scored by another doctoral student in clinical psychology. Two-way, random effects, absolute agreement ICCs (A,1; McGraw & Wong, 1996) were excellent (see Table 1).

Holaday and Sparks (2001) developed a set of comprehensive scoring guidelines for the MOA that largely follow Urist's (1977) initial article but expand on scoring guidelines and examples for each scoring level. Holaday and Sparks permitted reliable scoring even among raters naïve to OR theory. Using these guidelines, G. Bombel scored all Rorschach protocols and calculated the four index scores (MOAx, MOAb, MOAp, and MOAHI). The MOA Scale was scored blind to the criterion variables in this study (i.e., the SCORS and other Rorschach variables) and patient diagnosis. To determine interrater reliability for the MOA Scale, we randomly selected 20 protocols and J. L. Milhura independently scored them. Agreement was excellent ($\kappa = .94$) regarding whether a Rorschach response was scorable with the MOA Scale (Cicchetti, 1994). Interrater reliability ICCs (A,1) for the MOA Scale indexes were also in the excellent range (see Table 1).

G. Bombel scored all TAT protocols (cards 1, 2, 3BM, 4, 10, and 13MF) with the CHR, AT, and EIR scales. When scoring with a particular SCORS scale, G. Bombel was unaware of participants' scores on other SCORS scales, MOA Scale scores, and patient diagnosis. For each participant, final scores for each scale were simple averages across the six stories. Another doctoral student in clinical psychology randomly selected and independently scored 20 protocols for the three SCORS scales. G. Bombel and this student trained to the "gold standard" story examples in the SCORS and SCORS–G manuals. Interrater reliability ICCs (A,1) were computed on the final scores for each scale and were in the fair to good range (Cicchetti, 1994; see Table 1).

The $r_{contrast\text{-}CV}$ procedure. In this study, the initial step in Westen and Rosenthal's (2003) procedure was to create predicted correlations between the MOA Scale and the criterion variables. Each of us first read brief descriptions of the MOA Scale, the criterion variables, and the concepts of psychopathology and object relations theory. Next, we independently predicted what the magnitude and direction of the observed correlation might be between the MOA Scale and each of the criterion variables while taking into account method variance. We did this twice. First, we made ratings assuming that the MOA Scale is primarily a measure of OR. For example, a moderate negative correlation might be predicted between the MOA Scale and CHR because they would appear to measure similar constructs given our assumption about the MOA Scale. The predicted correlation would be negative because lower MOA Scale scores, and higher CHR scores, represent healthier representations. We obtained the final MOA Scale as OR prediction set of correlations by averaging the predicted correlations across the three raters for each criterion variable. The average measures ICC (A,3; McGraw & Wong, 1996) among us was .96. Subsequently, we repeated these steps assuming that the MOA

Scale is primarily a measure of psychopathology. We obtained the final MOA Scale as pathology prediction set of correlations as before, and agreement among the raters was again excellent (ICC [A,3] = .94; Cicchetti, 1994).

Next, for the MOA Scale as OR prediction set, we generated exact lambda weights by subtracting the mean of the predicted correlations from each individual correlation and then squaring and summing these values. We obtained actual correlations between one of the MOA scales and the criterion variables, transformed them into Fisher's Z scores, and then multiplied them by the squared lambda weights for the predicted correlations. We used the sum of this column ($\sum \lambda Z_r$), along with the median intercorrelation (r_x) among the criterion variables, and the average squared observed correlation to compute $Z_{contrast}$(see Westen & Rosenthal, 2003, for more detailed procedures). We then obtained the p and t values associated with the $Z_{contrast}$ with the Distribution function in SPSS (SPSS Inc., Chicago, IL), and we computed $r_{contrast\text{ - }CV}$ as the square root of the squared $t$score divided by the sum of the squared $t$score and the degrees of freedom. We subsequently repeated the procedure using the MOA Scale as psychopathology severity prediction set. We then repeated the whole process for each of the MOA Scale indexes. Thus, for each MOA index, we generated two $r_{contrast\text{ - }CV}$ effect sizes, one assuming that the MOA Scale measures OR and one assuming that it measures psychopathology severity. Subsequently, we tested the differences between the OR and psychopathology effect sizes for each index for significance using a procedure for comparing two sets of contrasts (i.e., theories; Rosenthal, Rosnow, & Rubin, 2000). Finally, to identify which index was the better measure of OR and which was the better measure of psychopathology, we examined $r_{contrast\text{ - }CV}$ effect sizes and their 95% confidence intervals across indexes for each prediction set. Procedures have not been developed to compare such effect sizes statistically (R. Rosenthal, personal communication, April 13, 2008).

Results

Table 1 presents descriptive statistics for the MOA Scale indexes and criterion variables. The distribution of the WSum6 variable was skewed (2.72) and kurtotic (9.96) because 1 participant had an extremely high value; it was therefore transformed with a square root function (revised $M = 4.22$; $SD = 2.32$; skew = 0.913; kurtosis = 1.55). The observed correlations between the MOA Scale indexes and the criterion variables, as well as the set of competing predicted correlations, are in Table 2. The primary $r_{contrast\text{ - }CV}$ procedure data are presented in Table 3. As can be seen from the first row in Table 3, for all four indexes, the MOA Scale as pathology measure effect size was larger than the MOA Scale as OR measure effect size. The difference approached significance for MOAHI ($p = .07$), and reached significance for MOAx ($p = .03$). This latter index appears to be more sensitive to pathology than to OR, whereas the other indexes appear to be equally sensitive to both constructs. Following Cohen's (1992) guidelines, for both the MOA Scale as OR measure and MOA Scale as pathology measure, $r_{contrast\text{ - }CV}$ effect sizes were large (>.50) for the MOAHI, medium to large (.39–.49) for MOAx and MOAp, and small or negligible (≤.10) for MOAb.

For each of the prediction sets (OR, pathology), the 95% confidence interval (CI) around MOAb's $r_{contrast\text{ - }CV}$ effect size did not overlap the MOAHI and MOAx CIs. This suggests that these latter indexes are more sensitive as overall OR and pathology measures than MOAb is. CIs spanned about .30 points ($Zr \pm 1.96/\sqrt{N-3}$), regardless of prediction set, and all other CIs overlapped. However, overlapping CIs do not necessarily prevent mean differences from being significant (Wolfe & Hanley, 2002). As noted, the difference between MOAx OR and MOAx pathology $r_{contrast\text{ - }CV}$ effect sizes was significant ($p = .03$) despite an absolute value of .07 and overlapping CIs. MOAHI OR and pathology effect sizes differed from their MOAx and MOAp counterparts by .10 to .18 points, whereas MOAp effect sizes were about .30 points higher than corresponding MOAb effect sizes.

A surprising finding was that the effect sizes for the MOAb were negligible. In this sample, most participants had a MOAb

TABLE 2.—Predicted and observed correlations between the MOA Scale and criterion variables.

Criterion	Predicted Correlations[a]		Observed Correlations			
Variable	OR	Pathology	MOAx[b]	MOAb[b]	MOAp[b]	MOAHI[c]
CHR[c]	−.43	−.15	.32**	.08	.35**	−.27**
AT[c]	−.32	−.20	−.05	.18	−.21*	.12
EIR[c]	−.40	−.15	.02	−.06	.04	.06
GHR[c]	−.38	−.22	−.13	−.33**	.12	.17
PHR[b]	.46	.30	.28**	−.15	.42**	−.33**
HRV[c]	−.33	−.22	−.32**	−.01	−.34**	.38**
X - %[b]	.15	.32	.10	.10	.04	−.16
WDA%[c]	−.17	−.30	−.09	−.13	.00	.11
WSum6[b]	.23	.45	.25*	−.19	.33**	−.19
PTI[b]	.18	.45	.19	−.04	.25*	−.21*
S-CON[b]	.15	.43	.32**	−.02	.29**	−.33**
DEPI[b]	.15	.42	.21*	.12	.20*	−.32**
CDI[b]	.18	.37	.05	.34**	−.22*	−.09
CESI[b]	.18	.45	.30**	−.13	.26**	−.25*

Note. All analyses, $N = 100$. MOA = Mutuality of Autonomy Scale; OR = object relations: predicted correlations for the MOA Scale as an OR measure; Pathology = predicted correlations for the MOA Scale as a measure of psychopathology; MOAx = MOA Scale mean; MOAb = MOA Scale best score index; MOAp = Moa Scale pathology index; MOAHI = MOA Scale Health index; SCORS = Social Cognition and Object Relations Scale; CHR = Complexity of Human Representation SCORS scale; AT = Affect Tone SCORS scale; EIR = Emotional Investment in Human Relationships SCORS scale; PTI = Perceptual Thinking index; S-CON = Suicide Constellation; DEPI = Depression index; CDI = Coping Deficit index; CESI = Conceptual Ego Strength index.

[a] Signs of these correlations should be reversed when considering the MOAHI, which is reverse scored relative to the other MOA scales. [b] Lower scores are healthier. [c] Higher scores are healthier.

*$p < .05$, two-tailed. **$p < .01$, two-tailed.

TABLE 3.—MOA Scale $r_{contrast\text{ - construct validity (CV)}}$ effect size statistics for competing prediction sets across indexes.

	MOAx		MOAb		MOAp		MOAHI	
Quantity	OR	Pathology	OR	Pathology	OR	Pathology	OR	Pathology
$r_{contrast\text{ - }CV}$	.42	.49	.03	.10	.39	.41	.54	.59
95% CI								
From:	0.25	0.32	−0.17	−0.01	0.21	0.24	0.38	0.45
To:	0.57	0.63	0.23	0.30	0.54	0.56	0.66	0.71
$Z_{contrast}$	4.40	5.19	0.28	.97	4.01	4.21	−5.79	−6.50
$t_{contrast}$	4.64	5.58	0.28	.97	4.19	4.42	6.34	7.29
$p_{contrast}$	<.001	<.001	0.390	.166	<.001	<.001	<.001	<.001

Note. MOA = Mutuality of Autonomy Scale; OR = object relations: MOA Scale as OR measure; Pathology = MOA Scale as measure of psychopathology; MOAx = MOA Scale mean; MOAb = MOA Scale best score index; MOAp = MOA Scale pathology index; MOAHI = MOA Scale health index; CI = confidence interval.

score of 1 ($n = 52$) or 2 ($n = 42$); only 6 of them scored above a 2. Here, MOAb simply did not covary much with other variables, so relationships were small.

A closer examination of our SCORS data revealed some puzzling relationships between the SCORS variables and the MOA Scale indexes. The observed correlations between all four MOA indexes and the CHR scale were not in the expected direction; the same was true for the MOAx and MOAp scores with the EIR scale and for MOAb with the AT scale. This occurred despite the fact that the MOA scores showed an expected pattern of relationship with other criteria. These discrepancies led us to consider errors when scoring the SCORS scales and/or entering them into our database. To test this, we generated scatter plots between MOAx and the SCORS variables, identified outliers, and reexamined their scores on MOA and SCORS variables. Further, 20 participants' MOA and SCORS scores were chosen randomly and checked for data entry error. No scoring or data entry errors were identified. However, given the unexpected relationships, we reanalyzed the data without the SCORS variables (See Table 4).

The $r_{contrast\text{ - }CV}$ effect sizes for most indexes increased across both prediction sets when the SCORS variables were excluded; however, the increases were most pronounced for the OR predictions. The indexes demonstrated the same orders of magnitude described earlier, descending from MOAHI, MOAx, MOAp, to MOAb. This time, both MOAHI and MOAx effect sizes were large. MOAp and MOAb remained medium large and small, respectively, although both MOAb effect sizes reached significance in these analyses ($p \leq .03$). MOA Scale OR effect sizes were higher than pathology effect sizes for MOAHI and MOAp, and the reverse was true for MOAb; but ultimately, none of the OR–pathology effect size differences were significant.

We utilized Holaday and Sparks' (2001) revised MOA Scale scoring guidelines instead of Urist's (1977) original criteria. The reviewers of the initial version of this manuscript noted that these two sets of guidelines differ in how a response that portrays "Fighting between equals" would be scored. The former would assign a Level 2 based on the portrayed balance of power, whereas the latter would assign a Level 5 for fighting (lower scores are healthier). To address this issue, all 18 "Fighting between equals" responses were rescored as Level 5 (as per Urist, 1977), and the analyses were re-run. Differences between the new and initial effect sizes were negligible (.01–.04) and did not favor one set of criteria systematically. Therefore, we do not report these data in a table.

TABLE 4.—Mutuality of Autonomy (MOA) Scale $r_{contrast\text{ - }construct\ validity\ (CV)}$ effect size statistics for competing prediction sets across indexes without Social Cognition and Object Relations Scale data.

	MOAx		MOAb		MOAp		MOAHI	
Quantity	OR	Pathology	OR	Pathology	OR	Pathology	OR	Pathology
$r_{contrast\text{ - }CV}$	.56	.56	.18	.22	.47	.40	.62	.61
95% CI								
From:	0.41	0.41	−0.02	0.02	0.30	0.22	0.48	0.47
To:	0.68	0.68	0.36	0.40	0.61	0.56	0.73	0.72
$Z_{contrast}$	6.04	6.07	1.77	2.21	4.89	4.17	−6.90	−6.78
$t_{contrast}$	6.67	6.71	1.79	2.24	5.22	4.37	7.85	7.68
$p_{contrast}$	<.001	<.001	.038	.014	<.001	<.001	<.001	<.001

Note. OR = object relations: MOA Scale as OR measure; Pathology = MOA Scale as measure of psychopathology; MOAx = MOA Scale mean; MOAb = MOA Scale best score index; MOAp = MOA Scale Pathology index; MOAHI = MOA Scale Health index; CI = confidence interval.

DISCUSSION

In this study, we examined the construct validity of the MOA Scale using the $r_{contrast\text{ - }CV}$ procedure developed by Westen and Rosenthal (2003). We generated competing hypotheses—that the MOA Scale primarily measures OR and that it primarily measures psychopathology. Each hypothesis was represented as a pattern of predicted correlations between the MOA Scale and a set of criterion variables, and each was tested against the actual correlations in our sample. We then generated the $r_{contrast\text{ - }CV}$ effect sizes representing the degree of correspondence between each of these prediction sets (OR, psychopathology) and the actual relationships. In light of previous research findings, our initial round of analyses produced some puzzling relationships among the SCORS variables and between SCORS variables and some MOA Scale indexes. We therefore decided to repeat the analyses without the SCORS data. Magnitudes of the resulting $r_{contrast\text{ - }CV}$ effect sizes were large for the MOAHI, MOAx, and MOAp indexes and small for the MOAb index, although all were significant ($p < .05$). The differences between the MOA Scale as OR measure and MOA Scale as pathology measure $r_{contrast\text{ - }CV}$ effect sizes were not significant for any index; indeed, none of the effect sizes differed by more than .07 points. Thus, neither of the competing hypotheses was supported. The MOA Scale appears to be a good measure of OR quality as well as psychopathology severity, but it does not appear to discriminate between the two constructs.

Taken separately, our observed OR and pathology effect sizes are therefore consistent with studies that have supported the construct validity of the MOA Scale as a measure of object relations (Ackerman et al., 2001; Blatt et al., 1990; Fowler, Hilsenroth, & Handler, 1996; Urist, 1977) as well as studies in which MOA Scale scores have converged with indicators of psychopathology, such as diagnosis severity and lifetime psychosis severity (Harder et al., 1984). Furthermore, our data are consistent with a number of other studies that have shown that the MOA Scale is a significant predictor of behavior and an effective tool to discriminate among a variety of groups (Ackerman et al., 2000; Brown-Cheatham, 1993; Cook et al., 1995; Fowler et al., 2004; Fowler, Hilsenroth, & Nolan, 2000; Goddard & Tuber, 1989; Kavanagh, 1985; Leichsenring, 2004; Leifer, Shapiro, Martone, & Kassem, 1991; Sayler, Holmstrom, & Noshpitz, 1991; Strauss & Ryan, 1987; Tuber, 1983; Tuber, Frank, & Santostefano, 1989).

Although preliminary, our findings raise some practical issues. Classical psychoanalytic drive/structure theory holds that internal representations of self and others develop from the interplay of instincts, social contact with caregivers, and inherited characteristics such as temperament (Berg et al., 1993; Freud, 1949). In contrast, OR theory states that self-representations and other representations develop concomitant with, or even developmentally precede, other aspects of psychological structure such as the instincts and ego functions (Christopher, Bickhard, & Lambeth, 2001; Kernberg, 1986; Masterson, 1981). In either theory, the development of object representations necessarily covaries with the development of the personality, including into pathological outcomes. Given this theoretical

relationship, OR quality and psychopathology should be fundamentally intertwined and correlated constructs. Indeed, the relationship between our OR and pathology prediction sets in the second and third column of Table 2 reflected this ($r = .89$; ICC [A,1] = .78).

However, there has been an ongoing interest in developing assessment approaches or instruments that can home in on specific aspects of dynamic personality structure and functioning, for example, primary process phenomena (Holt, 2002); ego impairment (Perry & Viglione, 1991); defensive functioning (Lerner, 1990); core conflict (Luborsky & Crits-Cristoph, 1990); attachment (Main, Kaplan, & Cassidy, 1985); and of course, object representations (Urist, 1977), to name a few. When such measures are designed for use with the Rorschach, scoring guidelines focus on thematic and/or structural elements of response contents that relate theoretically to the constructs being assessed. For example, an MOA Scale score (level) for a particular response is assigned based on the OR developmental theme of any interaction portrayed. This scoring focus is an attempt to make the measure more sensitive to the target construct than to all other constructs (i.e., discrimination), even ones that might overlap theoretically. If the guidelines can accomplish this, the final score communicates information primarily about the target construct. If not, the score will communicate information about the target construct as well as other constructs the guidelines cannot "tune out." Here, the aggregated score would represent some unknowable proportion of information about target and nontarget constructs.

In this study, we examined precisely this capacity in the MOA Scale and found that it did not discriminate OR phenomena from general psychopathology. Consequently, interpreting MOA Scale index scores as relatively pure indicators of internalized object representation quality that are distinct from overall psychopathology would not be sound. At a more molecular level of interpretation, however, some authors have suggested that levels at the healthy end of the scale are generally more sensitive to OR phenomena than to psychopathology, whereas Levels 5, 6, and 7 are more sensitive to psychopathology (Berg et al., 1993; Blatt et al., 1990). Our data are consistent with this. Only two criterion variables correlated significantly with MOAb: GHR ($r = -.33$, $p < .01$) and the CDI ($r = .34$, $p < .01$). MOAp correlated significantly with S-CON ($r = .29$, $p < .01$) and DEPI ($r = .20$, $p < .05$), whereas WSum6 ($r = .33$, $p < .01$) and PTI ($r = .25$, $p < .05$) correlated better with MOAp than with any other index. PHR, which is scored from structural elements of responses, related better to MOAp ($r = .42$, $p < .01$) than did either of these two thought disorder variables.

The MOA Scale may be most clinically useful when interpretation begins with a careful review of response content and considers a range of additional scores to refine interpretation. This process can reveal the overall tone of the object representations in the protocol and clarify how OR and pathology phenomena contributed to final index scores (Berg et al., 1993). For example, an MOAx score of 5.32 might reflect OR quality more than psychopathology if X – % = .15, PTI = 2, and S-CON = 4. On the other hand, psychopathology would likely contribute more to this same MOAx score if it were accompanied by X – % = .35, PTI = 4, and S-CON = 7.

One possible broader implication of these results is that theme-based Rorschach approaches to assessing OR phenomena, in general, are inherently unable to discriminate OR phenomena and psychopathology at the aggregated index score level of analysis. Indeed, OR quality may be too complicated to assess adequately in a single Rorschach modality. Measures that focus almost exclusively on theme (Coonerty, 1986; Kwawer, 1979; Mayman, 1967; Urist, 1977) or structural aspects of responses (Blatt et al., 1976; Pruitt & Spilka, 1964) may not be sampling the full breadth of how this construct is expressed on the Rorschach. Here, it could be that OR phenomena and psychopathology are distinct enough to assess more purely as separate constructs but that Rorschach measures are only sensitive to a fraction of the actual phenomena that would discriminate between the two. Alternately, the two constructs are simply not distinguishable enough to measure more purely than we presently do. Our data may support the latter hypothesis. As noted earlier, based on the two sets of predicted correlations, the independent judges expected that the MOA Scale as a pure measure of OR would behave much like the MOA Scale as a pure measure of pathology ($r = .89$; ICC [A,1] = .78). Importantly, these judgments were reliable, with average measures ICCs across judges for the OR and pathology prediction sets of .96 and .94, respectively. Still, to further clarify the issue, it would be helpful to use Westen and Rosenthal's (2003) $r_{contrast\text{ - }CV}$ procedure to examine implicit and performance OR measures with samples that are representative of the normal population.

In our sample ($N = 100$; Total $r = 2{,}271$), 18 responses across 13 protocols portrayed "Fighting between equals." Scoring these responses as Level 5 (Urist, 1977) or Level 2 (Holaday & Sparks, 2001) ultimately made little difference. The $r_{contrast\text{ - }CV}$ effect sizes generated from the Urist (1977) and Holaday and Sparks (2001) scoring guidelines were essentially equivalent. This suggests that the "Fighting between equals" response occurs infrequently enough that Holaday and Sparks' (2001) MOA Scale scoring guidelines can be used in lieu of Urist's (1977) in research and clinical settings. This is fortunate because the Holaday and Sparks criteria are more comprehensive and provide a great deal more guidance about scoring decisions than the Urist guidelines. Furthermore, as noted, Urist's (1977; Urist & Shill, 1982) published descriptions of MOA Scale are inconsistent across Levels 2 through 6, which can create scoring confusion.

MOAx is one of the most reported MOA indexes in the literature (Ackerman et al., 2000; Blatt et al., 1990; Fowler et al., 2004; Goddard & Tuber, 1989; Leifer et al., 1991; Mazor, Alfa, & Gampel, 1993; Ryan et al., 1985; Spear & Sugarman, 1984; Strauss & Ryan, 1987; Tuber, 1989; Tuber & Coates, 1989), and it has frequently been used as the sole MOA Scale index (Canetto, Feldman, & Lupei, 1989; Donahue & Tuber, 1993; Fowler et al., 1995; Harder et al., 1984; Hart & Hilton, 1988; J. F. Murray, 1985; Tuber et al., 1989). Results of this study, however, suggest that it may not be the best MOA Scale index for assessing its target constructs, OR quality and psychopathology. MOAHI, which is computed as a ratio of healthy to primitive scores relative to the number of scorable MOA responses, appears to be the most sensitive MOA Scale index. MOAHI $r_{contrast\text{ - }CV}$ effect sizes were the largest of any index regardless of prediction set (OR or pathology) and whether the SCORS variables were included in the analyses.

A limitation of this study was the uncertainty regarding the validity of the CHR, AT, and EIR scales. The interrater reliability values ranged from fair to good and suggest that

these scales were scored with an acceptable degree of reliability by raters who previously documented reliable scoring relative to gold standard protocols; however, the correlations among these scales, as well as the relationships between these scales and some MOA Scale indexes, did not correspond well to previously published data. For example, Ackerman et al. (2001) obtained expected significant correlations between MOAb and two of the SCORS scales, AT and EIR, whereas we found small or negligible relationships between MOAb and all three SCORS scales. Part of this, as noted, may be due to the properties of the MOAb index score distribution in our sample. Our results also differed from Ackerman et al.'s (2001) in that we obtained significant relationships between MOAp and two SCORS scales, CHR and AT. The significant relationships between MOAp and CHR, and MOAHI and CHR, were in unexpected directions, although MOAp and MOAHI generally behaved as expected with other variables.

A second limitation of this study was the limited number of assessment methods used to obtain personality data. We used only implicit techniques. Arguably, however, the Rorschach and TAT tasks are different enough that placing them in the same category of assessment method is challenging; indeed, it may be inappropriate (Meyer & Kurtz, 2006). Still, the design of any future replication of this project would do well to include behavioral measures among the criterion variables.

ACKNOWLEDGMENTS

We thank the anonymous reviewers for their helpful comments on earlier versions of this article. We also thank Prachi Kene, Margaret Chan, Nick Katko, and Aaron Upton for their help with interrater reliability scoring.

REFERENCES

Ackerman, S. J., Hilsenroth, M. J., Clemence, A. J., Weatherill, R., & Fowler, J. C. (2000). The effects of social cognition and object representation on psychotherapy continuation. *Bulletin of the Menninger Clinic, 64,* 386–408.

Ackerman, S. J., Hilsenroth, M. J., Clemence, A. J., Weatherill, R., & Fowler, J. C. (2001). Convergent validity of Rorschach and TAT scales of object relations. *Journal of Personality Assessment, 77,* 295–306.

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text revision). Washington, DC: Author.

Barron, F. (1953). An ego-strength scale which predicts response to psychotherapy. *Journal of Consulting Psychology, 17,* 327–333.

Bell, M. (1991). *An introduction to the Bell Object Relations Reality Testing Inventory*. (Available from Morris Bell, Psychology Service, 116B, Veterans Affairs Medical Center, West Haven, CT 06516).

Berg, J. L., Packer, A., & Nunno, V. J. (1993). A Rorschach analysis: Parallel disturbance in thought and in self/object representation. *Journal of Personality Assessment, 61,* 311–323.

Bessell, H. (1984). *The love test*. New York: William Morrow.

Blais, M. A., Hilsenroth, M. J., Fowler, J. C., & Conboy, C. A. (1999). A Rorschach exploration of the *DSM–IV* borderline personality disorder. *Journal of Clinical Psychology, 55,* 563–572.

Blatt, S. J., Brenneis, C. B., Schimek, J. G., & Glick, M. (1976). Normal development and psychopathological impairment of the concept of the object on the Rorschach. *Journal of Abnormal Psychology, 85,* 364–373.

Blatt, S. J., & Ritzler, B. A. (1974). Thought disturbance and boundary disturbances in psychosis. *Journal of Consulting and Clinical Psychology, 42,* 370–381.

Blatt, S. J., Tuber, S. B., & Auerbach, J. S. (1990). Representation of interpersonal interactions on the Rorschach and level of psychopathology. *Journal of Personality Assessment, 54,* 711–728.

Bombel, G. A. (2006). *A meta-analysis of interrater scoring reliability for the Rorschach Mutuality of Autonomy (MOA) Scale*. Unpublished master's thesis, University of Toledo, Toledo, Ohio.

Brown-Cheatham, M. (1993). The Rorschach Mutuality of Autonomy Scale in the assessment of Black father-absent male children. *Journal of Personality Assessment, 61,* 524–530.

Burns, B., & Viglione, D. J. (1996). The Rorschach Human Experience variable, interpersonal relatedness, and object representations in nonpatients. *Psychological Assessment, 8,* 92–99.

Burke, W. F., Friedman, G., & Gorlitz, P. (1988). The psychoanalytic Rorschach profile: An integration of drive, ego, and object relations perspectives. *Psychoanalytic Psychology, 5,* 193–212.

Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). *MMPI–2: Minnesota Multiphasic Personality Inventory–2: Manual for administration and scoring.* Minneapolis: University of Minnesota Press.

Canetto, S. S., Feldman, L. B., & Lupei, R. L. (1989). Suicidal persons and their partners—Individual and interpersonal dynamics. *Suicide and Life-Threatening Behavior, 19*, 237–248.

Christopher, J. C., Bickhard, M. H., & Lambeth, G. S. (2001). Otto Kernberg's object relations theory: A metapsychological critique. *Theory & Psychology, 11,* 687–711.

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6,* 284–290.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155–159.

Cook, B., Blatt, S., & Ford, R. (1995). The prediction of therapeutic response to long-term intensive treatment of seriously disturbed young adult inpatients. *Psychotherapy Research, 5,* 218–230.

Coonerty, S. (1986). An exploration of separation-individuation themes in the borderline personality disorder. *Journal of Personality Assessment, 50,* 501–511.

Cooper, S. H., Perry, J. C., Hoke, L., & Richman, R. (1985). Transitional relatedness and borderline personality disorder. *Psychoanalytic Psychology, 2,* 115–128.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52,* 281–302.

Dean, K. L., Viglione, D. J., Perry, W., & Meyer, G. J. (2007). A method to optimize the response range while maintaining Rorschach Comprehensive System validity. *Journal of Personality Assessment, 89,* 149–161.

Dao, T. K., & Prevatt, F. (2006). A psychometric evaluation of the Rorschach Comprehensive System's Perceptual Thinking Index. *Journal of Personality Assessment, 86,* 180–189.

Donahue, P. J., & Tuber, S. B. (1993). Rorschach adaptive fantasy images and coping in children under severe environmental stress. *Journal of Personality Assessment, 60,* 421–434.

Exner, J. E. (1986). *The Rorschach: A Comprehensive System: Vol. 1. Basic foundations* (2nd ed.). New York: Wiley.

Exner, J. E. (1993). *The Rorschach: A Comprehensive System: Vol. 1. Basic foundations* (3rd ed.). Oxford, England: Wiley.

Exner, J. E. (2003). *The Rorschach: A Comprehensive System: Vol. 1. Basic foundations and principles of interpretation* (4th ed.). Hoboken, NJ: Wiley.

Fisher, P. H., Sperling, M. B., & Carr, A. C. (1990). Assessment of adult relatedness: A review of empirical findings from object relations and attachment theories. *Journal of Personality Assessment, 55,* 499–520.

Fowler, J. C., Ackerman, S. J., Speanburg, S., Bailey, A., Blagys, M., & Conklin, A. C. (2004). Personality and symptom change in treatment-refractory inpatients: Evaluation of the phase model of change using Rorschach, TAT, and *DSM–IV* Axis V. *Journal of Personality Assessment, 83,* 306–322.

Fowler, J. C., Addelson, B., & Clemence, A. J. (2006). (2006, March). *A meta-analysis of the Mutuality of Autonomy Scale.* Paper presented at the Annual Meeting of the Society for Personality Assessment, Washington, DC.

Fowler, C., Hilsenroth, M. J., & Handler, L. (1995). Early memories: An exploration of theoretically derived queries and their clinical utility. *Bulletin of the Menninger Clinic, 59,* 79–98.

Fowler, J. C., Hilsenroth, M. J., & Handler, L. (1996). A multimethod approach to assessing dependency: The early memory dependency probe. *Journal of Personality Assessment, 67,* 399–413.

Fowler, J. C., Hilsenroth, M. J., & Nolan, E. (2000). Exploring the inner world of self-mutilating borderline patients: A Rorschach investigation. *Bulletin of the Menninger Clinic, 64,* 365–385.

Fowler, J. C., Piers, C., Hilsenroth, M. J., Holdwick, D. J., & Padawer, J. R. (2001). The Rorschach Suicide Constellation: Assessing various degrees of lethality. *Journal of Personality Assessment, 76,* 333–351.

Frank, G. (1995). On the assessment of self and object representations from the Rorschach: A review of the research and commentary. *Psychological Reports, 76,* 659–671.

Freud, S. (1949). *An outline of psychoanalysis*. New York: Norton.

Goddard, R., & Tuber, S. (1989). Boyhood separation anxiety disorder: thought disorder and object relations psychopathology as manifested in Rorschach imagery. *Journal of Personality Assessment, 53,* 239–252.

Goldberg, L. R. (1965). Diagnosticians vs. diagnostic signs: The diagnosis of psychosis vs. neurosis from the MMPI–2. *Psychological Monographs, 79*(9, Whole No. 602).

Harder, D. W., Greenwald, D. F., Wechsler, S., & Ritzler, B. A. (1984). The Urist Rorschach Mutuality of Autonomy Scale as an indicator of psychopathology. *Journal of Clinical Psychology, 40,* 1078–1083.

Hart, B., & Hilton, I. (1988). Dimensions of personality organization as predictors of teenage pregnancy risk. *Journal of Personality Assessment, 52*, 116–132.

Hibbard, S., Hilsenroth, M. J., Hibbard, J. K., & Nash, M. R. (1995). A validity study of two projective object representations measures. *Psychological Assessment, 7,* 432–439.

Hilsenroth, M., Stein, M., & Pinsker, J. (2004). *Social cognition and object relations scale: Global rating method.* Unpublished manuscript, The Derner Institute of Advanced Psychological Studies, Adelphi University, Garden City, NJ.

Holaday, M., & Sparks, C. L. (2001). Revised guidelines for Urist's Mutuality of Autonomy Scale (MOA). *Assessment, 8,* 145–154.

Holt, R. R. (2002). Quantitative research on the primary process: method and findings. *Journal of the American Psychoanalytic Association, 50,* 457–482.

Huprich, S. K., & Greenberg, R. P. (2003). Advances in the assessment of object relations in the 1990s. *Clinical Psychology Review, 23,* 665–698.

Jørgenson, K., Andersen, T. J., & Dam, H. (2000). The diagnostic efficiency of the Rorschach Depression Index and the Schizophrenia Index: A review. *Assessment, 7,* 259–280.

Kavanagh, G. G. (1985). Changes in patients' object representations during psychoanalysis and psychoanalytic psychotherapy. *Bulletin of the Menninger Clinic, 49,* 546–564.

Kernberg, O. K. (1986). *Severe personality disorders: Psychotherapeutic strategies*. New Haven, CT: Yale University Press.

Kleiger, J. H. (1992). A conceptual critique of the EA:es comparison in the Comprehensive Rorschach System. *Psychological Assessment, 4,* 288–296.

Kwawer, J. S. (1979). Borderline phenomena, interpersonal relations, and the Rorschach test. *Bulletin of the Menninger Clinic, 43,* 515–524.

Leichsenring, F. (2004). Quality of depressive experiences in borderline personality disorders: Differences between patients with borderline personality disorder and patients with higher levels of personality organization. *Bulletin of the Menninger Clinic, 68,* 9–22.

Leifer, M., Shapiro, J. P., Martone, M. W., & Kassem, L. (1991). Rorschach assessment of psychological functioning in sexually abused girls. *Journal of Personality Assessment, 56,* 14–28.

Lerner, P. M. (1990). Rorschach assessment of primitive defenses: A review. *Journal of Personality Assessment, 54,* 30–46.

Luborsky, L., & Crits-Cristoph, P. (1990). *Understanding transference: The core conflictual relationship theme method*. New York: Basic Books.

Main, M., Kaplan, N., & Cassidy, J. (1985). Security in infancy, childhood, and adulthood: A move to the level of representation. In I. Bretherton & E. Waters (Eds.), *Monographs of the Society for Research in Child development, 50,* 66–104.

Masterson, J. F. (1976). *Psychotherapy of the borderline adult*. New York: Brunner/Mazel.

Masterson, J. F. (1981). *The narcissistic and borderline disorders: An integrated developmental approach*. New York: Brunner/Mazel.

Mayman, M. (1967). Object representations and object relationships in Rorschach responses. *Journal of Projective Techniques & Personality Assessment, 31,* 17–24.

Mazor, A., Alfa, A., & Gampel, Y. (1993). On the thin line between connection and separation: The individuation process, from cognitive and object relations perspectives, in Kibbutz adolescents. *Journal of Youth and Adolescents*, *22*, 641–669.

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods, 1,* 30–46.

McWilliams, N. (1994). *Psychoanalytic diagnosis: Understanding personality structure in the clinical process*. New York: Guilford.

Meyer, G. J. (1992). Response frequency problems in the Rorschach: Clinical and research implications with suggestions for the future. *Journal of Personality Assessment, 58,* 231–244.

Meyer, G. J., & Kurtz, J. E. (2006). Advancing personality assessment terminology: Time to retire "objective" and "projective" as personality test descriptors. *Journal of Personality Assessment, 87,* 223–225.

Meyer, G. J., & Resnick, J. (1996). *Assessing ego impairment: Do scoring procedures make a difference?* Paper presented at the XV International Congress of Rorschach and Projective Methods, Boston, MA.

Murray, H. A. (1943). *Thematic Apperception Test manual.* Cambridge, MA: Harvard University Press.

Murray, J. F. (1985). Borderline manifestations in the Rorschachs of male transsexuals. *Journal of Personality Assessment, 49*, 454–466.

Perry, W., & Viglione, D. J. (1991). The Ego Impairment Index as a predictor of outcome in melancholic depressed patients treated with tricyclic antidepressants. *Journal of Personality Assessment, 56,* 487–501.

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

Pruitt, W. A., & Spilka, B. (1964). Rorschach empathy-object relationship scale. *Journal of Projective Techniques & Personality Assessment, 28,* 311–336.

Resnick, J. (1994). *Rorschach assessment of ego functioning: A comparison of the EII and CESI*. Unpublished master's thesis, Loyola University of Chicago, Chicago, Illinois.

Resnick, J., & Meyer, G. J. (1995, March). *Rorschach assessment of ego functioning: A comparison of the EII and the CESI*. Paper presented at the annual meeting of the Society for Personality Assessment, Atlanta, GA.

Rorschach, H. (1942). *Psychodiagnostik* [Psychodiagnostics] (Han Huber, Trans.). Bern, Switzerland: Verlag. (Original work published 1921)

Rosenthal, R., Rosnow, R. L., & Rubin, D. B. (2000). *Contrasts and effect sizes in behavioral research: A correlational approach*. New York: Cambridge University Press.

Ryan, R. M., Avery, R. R., & Grolnick, W. S. (1985). A Rorschach assessment of children's mutuality of autonomy. *Journal of Personality Assessment, 49,* 6–12.

Salyer, K. M., Holmstrom, R. W., & Noshpitz, J. D. (1991). Learning disabilities as a childhood manifestation of severe psychopathology. *American Journal of Orthopsychiatry, 61,* 230–240.

Smith, T. E. (1993). Measurement of object relations: A review. *Journal of Psychotherapy Practice & Research, 2,* 19–37.

Smith, S. R., Baity, M. R., Knowles, E. S., & Hilsenroth, M. J. (2001). Assessment of disordered thinking in children and adolescents: The Rorschach Perceptual-Thinking Index. *Journal of Personality Assessment, 77,* 447–463.

Spear, W. E., & Sugarman, A. (1984). Dimensions of internalized object relations in borderline and schizophrenic patients. *Psychoanalytic Psychology, 1*, 113–129.

Strauss, J., & Ryan, R. M. (1987). Autonomy disturbances in subtypes of anorexia-nervosa. *Journal of Abnormal Psychology, 96,* 254–258.

Stricker, G., & Healey, B. J. (1990). Projective assessment of object relations: A review of the empirical literature. *Psychological Assessment, 2,* 219–230.

Tuber, S. B. (1983). Children's Rorschach scores as predictors of later adjustment. *Journal of Consulting & Clinical Psychology, 51,* 379–385.

Tuber, S. B. (1989). Assessment of children's object-representations with the Rorschach. *Bulletin of the Meninger Clinic, 53*, 432–441.

Tuber, S. (1992). Empirical and clinical assessments of children's object relations and object representations. *Journal of Personality Assessment, 58,* 179–197.

Tuber, S., & Coates, S. (1989). Indices of psychopathology in the Rorschachs of boys with severe gender identity disorder: A comparison with normal control subjects. *Journal of Personality Assessment, 53,* 100–112.

Tuber, S. B., Frank, M. A., & Santostefano, S. (1989). Children's anticipation of impending surgery: Shifts in object-representational paradigms. *Bulletin of the Menninger Clinic, 53,* 501–511.

Urist, J. (1977). The Rorschach Test and the assessment of object relations. *Journal of Personality Assessment, 41,* 3–9.

Urist, J., & Shill, M. (1982). Validity of the Rorschach Mutuality of Autonomy Scale: A replication using excerpted responses. *Journal of Personality Assessment, 46,* 450–454.

Viglione, D. J., Perry, W., Jansak, D., Meyer, G., & Exner, J. E. (2003). Modifying the Rorschach Human Experience variable to create the Human Representational variable. *Journal of Personality Assessment, 81,* 64–73.

Westen, D., Lohr, N., Silk, K., Kerber, K., & Goodrich, S. (1990). *Social cognition and object relations scale (SCORS): Manual for coding TAT data (5th ed.)*. Unpublished manuscript, University of Michigan, Ann Arbor.

Westen, D., & Rosenthal, R. (2003). Quantifying construct validity: Two simple measures. *Journal of Personality and Social Psychology, 84,* 608–618.

Wolfe, R., & Hanley, J. (2002). If we're so different, why do we keep overlapping? When 1 plus 1 doesn't make 2. *Canadian Medical Association Journal, 166,* 65–66.