

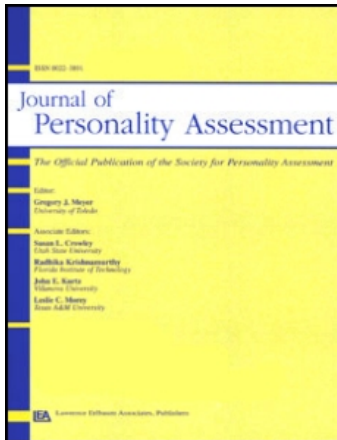
This article was downloaded by: [HJPA - Society of Personality Assessment Member]

On: 12 September 2008

Access details: Access Details: [subscription number 901432949]

Publisher Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



## Journal of Personality Assessment

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title-content=t775653663>

### The Rorschach in the Context of Performance-Based Personality Assessment

Robert E. McGrath<sup>a</sup>

<sup>a</sup> School of Psychology, Farleigh Dickinson University,

Online Publication Date: 01 September 2008

**To cite this Article** McGrath, Robert E.(2008)'The Rorschach in the Context of Performance-Based Personality Assessment',Journal of Personality Assessment,90:5,465 — 475

**To link to this Article:** DOI: 10.1080/00223890802248760

**URL:** <http://dx.doi.org/10.1080/00223890802248760>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

# The Rorschach in the Context of Performance-Based Personality Assessment

ROBERT E. MCGRATH

*School of Psychology, Fairleigh Dickinson University*

In this article, I draw on literature concerning the current status of the Rorschach Inkblot Method (Exner, 2002), performance-based personality measurement, direct versus indirect personality assessment, and dual-process models to suggest a reconceptualization of the Rorschach. The goal is to offer an updated conceptual framework for thinking about the Rorschach that can potentially be used both to enhance the image of the method and to provide direction for future research on the validity of specific Rorschach scores.

The current status of the Rorschach Inkblot Method, and the associated Comprehensive System used for its scoring and interpretation (Exner, 2002), was probably best summarized by Hunsley and Bailey (1999) when they stated, “the Rorschach has the dubious distinction of being, simultaneously, the most cherished and the most reviled of all psychological assessment tools” (p. 266). On one hand, opponents have presented wide-ranging criticisms of the method (e.g., Garb, Wood, Lilienfeld, & Nezworski, 2005; Lilienfeld, Wood, & Garb, 2000; Wood, Nezworski, & Lilienfeld, 2003; Wood, Nezworski, & Stejskal, 1996). At the same time, the Rorschach remains among the most commonly used instruments in clinical settings (e.g., Archer & Newsom, 2000; Camara, Nathan, & Puente, 2000); and its advocates are often passionate in its defense (e.g., Viglione & Hilsenroth, 2001; Weiner, 2001). Buttressing its continued use is the perception among many clinicians that the Rorschach can reveal information unavailable through other popular assessment techniques (Weiner, 1999). Furthermore, even critics of the Rorschach admit it is a valid indicator of certain clinical constructs including thought disorder and treatment prognosis and may be a valid predictor of others as well (Garb et al., 2005).

One particularly important point raised by the critics of the Comprehensive System (Garb et al., 2005) is the number of variables that still have not been adequately validated even though the system has been in use for over 20 years. The list includes many of the system’s key variables, the D score providing just one example. The well-established bias toward the publication of significant outcomes (Dickersin, 2005) would suggest that a variable for which no evidence of validity has been published is probably a variable of at best modest validity. In other cases, the validity evidence is minimal, consisting of at most one to two studies, and in such cases, the concern is that the published literature is devoted to those studies that have generated the largest effect size estimates for those variables (Schmidt, 1996). The number of Rorschach variables that have been inadequately validated suggests that global meta-analyses of Rorschach effects

(e.g., Hiller, Rosenthal, Bornstein, Berry, & Brunell-Neuleib, 1999; Parker, Hanson, & Hunsley, 1988), even if conducted in a methodological rigorous manner, may tend to overestimate the validity of Rorschach variables as a whole. Furthermore, it is possible Exner’s (2002) decision to build his system on codes derived from the five earlier systems of Rorschach scoring and interpretation, as opposed to focusing on those Rorschach variables with the strongest validity evidence regardless of their source, may have resulted in a less globally valid system than was possible.

The universe of variables that has been created for scoring from a Rorschach protocol is substantial: The Comprehensive System alone consists of dozens of frequencies, ratios, and proportions. Many other variables are available outside the system. An essential question for an efficient research program for creating an optimally valid approach to Rorschach scoring and interpretation has to do with which scores should be emphasized in future research. This question in part serves as motivation for this article. I use several bodies of research not traditionally applied to the Rorschach to suggest a conceptual framework for the Rorschach that provides guidelines for future research on the construct validity of Rorschach scores. To begin, I discuss recent efforts to broaden the context in which the Rorschach is framed.

## PERFORMANCE-BASED PERSONALITY ASSESSMENT

For many years, the Rorschach has been widely perceived as a *projective* instrument. This term has been used to indicate instruments that require responding to ambiguous stimuli based on the assumption, derived from psychoanalytic theory, that responses to such stimuli are determined primarily by the projection of internal wishes or conflicts. Recently, supporters of the Rorschach have questioned this assumption, beginning with Exner (1989). Although he acknowledged that the psychoanalytic process of projection can play a role in the identification or description of a percept, particularly when a percept is unusual or grossly inconsistent with the blot, Exner (1989) argued that the task set for the respondent by the question “what might this be?” is primarily perceptual. More recently, the *Journal of Personality Assessment* has rejected the use of the term *projective* as a label for personality indicators (Meyer & Kurtz, 2006), and Kubiszyn

Received October 19, 2007; Revised January 17, 2008.  
Address correspondence to Robert E. McGrath, School of Psychology, Fairleigh Dickinson University, Teaneck-Hackensack Campus, 1000 River Road T-WHI-01, Teaneck, NJ 07666; Email: mcgrath@fd.edu

et al. (2000) recommended the terminology *performance-based personality test* instead.<sup>1</sup>

This alternative connotes a more descriptive and theory-neutral perspective on the Rorschach. It also permits comparisons between the Rorschach and instruments with very different historical roots than the ambiguous-stimulus instruments traditionally considered projective. I use these comparisons to develop some hypotheses about the best future course for the development of the inkblot method.

As a starting point, it will be helpful to define the parameters of performance-based personality tests. The reference to personality is presumably intended to distinguish the instruments under consideration from performance-based indicators used to evaluate "abilities," such as intellectual, psychoeducational, neuropsychological, and competency tests. Personality can be viewed as the totality of the individual's characteristic style of responding to the environment distinguished from abilities in particular by the consideration of both cognitive and affective factors (Block, 2002; Mayer, 2005). That is, personality exceeds abilities by encompassing values, beliefs, preferences, and other affectively tinged components of characteristic functioning, a domain that also includes some of the more stable elements of psychopathology (e.g., Millon, Millon, Meagher, Grossman, & Ramnath, 2004). Given the nascent state of neuroscience, the degree to which abilities and personality reflect distinct aspects of brain function remains an open question. What is clear is that psychologists find the distinction between what a person can do and who a person is to be a useful one.

The other term in the phrase, "performance-based," is intended to distinguish between measures of ability and personality that involve behavioral sampling as opposed to other instruments relevant to personality such as the Minnesota Multiphasic Personality Inventory (MMPI; Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989) that focus on self-description by the respondent. Although typically more difficult to administer than self-report measures, the use of performance-based measures is considered justified by their proponents because they are used to gauge a person's style of responding to the instrument, under the assumption that this style of responding indicates something about the person's approach to environmental events outside the testing situation.

#### THE RORSCHACH AND THE IMPLICIT ASSOCIATION TEST

Reframing the Rorschach as a performance-based personality indicator suggests a similarity with any measurement device that uses the individual's performance on a series of tasks as a means of learning about that individual's style of functioning in affectively tinged situations. Another example of this class, and the one that is the primary focus of subsequent comparisons, is the Implicit Association Test (IAT; Greenwald, McGhee, &

Schwartz, 1998). The IAT is a computer-administered instrument that uses reaction times as the basis for inferences about the relative strength of associations. For example, the IAT might be used to present four types of stimuli in random order: positive words, negative words, pictures of flowers, and pictures of snakes. In one block of trials, respondents are instructed to hit the D key when they see either a positive word or a picture of a flower, and the K key when they see a negative word or a snake. In another block, the associations between stimuli are reversed, so D is used for positive words and snakes, K for negative words and flowers. Additional blocks are used to counterbalance the keys. Most individuals are expected to respond more quickly in the first block, where flowers are associated with positive words and snakes with negative words, than in the block where the association is reversed because cultural factors tend to encourage the association of flowers with positive and snakes with negative valuations.

The IAT is a method of data gathering rather than a specific instrument. The same has been said of the Rorschach (Weiner, 1994) because the latter is inherently a source of data rather than a consistent approach to quantifying that data, but because the set of stimuli used is at the discretion of the researcher, the claim is even truer for the IAT. Although used at first primarily to study social attitudes, which are affectively tinged but not generally treated as a component of personality because they are specific to certain stimulus objects in the person's environment, the IAT methodology has since been used in connection with a variety of clinical and personality constructs. For example, Nock and Banaji (2007) found the strength of association between the concepts of self-injury and self-reference was an independent predictor of level of suicide risk in adolescents. Other research has looked at such topics as associations with alcohol among patients undergoing rehabilitation for alcoholism (De Houwer, Crombez, Koster, & De Beul, 2004) and the relationship between extraversion and self-referential versus other-referential words (Schmukle & Egloff, 2005).

The IAT is one of a variety of performance-based measurement methods to emerge out of the social-personality literature in recent years (De Houwer & Moors, 2007). Still other performance-based tests have been developed with more strictly clinical goals in mind. Several tests have become popular in the evaluation of attention deficit hyperactivity disorder, including the Conners (1994) Continuous Performance Test (CPT) and the Test of Variables of Attention (TOVA; Greenberg, 1990). These instruments involve sequential presentation of stimuli in random order, with instructions to respond to one stimulus differently than all others. Variables such as response time and errors of omission and commission are used as indicators of hyperactivity, inattention, and impulsivity.

The CPT and TOVA are particularly interesting because they blur the distinction between abilities and personality. How accurately a person with attention deficit hyperactivity disorder performs on the task is believed to be reflective of an important influence on their personal and social identity. The IAT is the primary focus in my subsequent discussion of recent performance-based measures, however, because it is by far the most widely studied: Over 200 studies have been published using the IAT since its introduction 9 years ago. It is worth noting that many of the comments to be made about the IAT apply to other performance-based tests of personality functioning, just as many of the conclusions to be drawn about the Rorschach can

<sup>1</sup>Bornstein (2007b) and Schultheiss (2007) have suggested alternative classification models for psychological measures in response to Meyer and Kurtz (2006). The concept of performance-based personality tests used in this article encompasses four categories of instruments listed in Bornstein's (2007b) Table 1: stimulus attribution, performance based, constructive, and observational. It also parallels Schultheiss's category of nondeclarative measures. However, Schultheiss's association of nondeclarative measures with specific mental processes and with unconscious activity may be too restrictive, as my later discussion suggests.

be applied to other personality instruments that use ambiguous stimuli.

To reinforce the association between the Rorschach and the IAT, several examples can be provided of instances in which researchers working with the two instruments have struggled with the same issues while using different terminology. Prior literature on personality instruments presenting ambiguous stimuli has suggested performance-based indicators differ from self-report in that the former focus on implicit motivations, whereas the latter have to do with self-attributed motivations (McClelland, Koestner, & Weinberger, 1989). The social/personality literature has tended to distinguish instead between *implicit* and *explicit* measurement (Greenwald & Banaji, 1995). The latter terminology was derived from the parallel distinction drawn between implicit and explicit memory functioning (Roediger, 1990) and reflects the closer connection between the social/personality and experimental literatures on mental activity. There is also a literature that has suggested the terminology *mental processes* versus *mental experiences* (e.g., Nosek, 2007). This last alternative has the advantage of avoiding the implication the individual is unaware of “implicit” activity such as automatic attitudes, an implication that has been questioned in a number of studies (e.g., Fazio & Olson, 2003; Gawronski, Hofmann, & Wilbur, 2006; Olson, Fazio, & Hermann, 2007). The distinction between mental processes (gauged via performance-based indicators) and mental experiences (gauged via self-report indicators) provides a framework with broad implications for the understanding of all personality measurement, whether derived from the clinical or social/experimental literature. Although use of the terms implicit and explicit is widespread in the literature, including in the title of the IAT, I use the adjectives *automatic* and *deliberate* in subsequent discussion to avoid any implications about level of awareness.

A second parallel emerges from research that has compared outcomes on an automatic indicator of mental process with a self-report indicator of mental experience, for example, correlating an IAT and self-report measure of self-esteem or the Rorschach Depression Index with depression scales from the MMPI. These correlations have proven to be small on average (Archer & Krishnamurthy, 1993a, 1993b; Hofmann, Gawronski, Gschwendner, Le, & Schmitt, 2005; Nosek, 2005). For example, Poehlman, Uhlmann, Greenwald, and Banaji (2005) found nine studies that had compared self-report to IAT measures of clinical issues. The weighted average of these correlations was .18. Similar comparisons involving the Rorschach and MMPI produced a mean  $r$  across 17 analyses of .06 (Meyer, Riethmiller, Brooks, Benoit, & Handler, 2000). Various attempts have been made to understand the reasons for this lack of convergence (e.g., Bornstein, 2002; Gawronski, LeBel, & Peters, 2007; Meyer et al., 2000; Olson et al., 2007), although currently the IAT and Rorschach literatures on this topic are evolving in complete isolation from each other.

A third parallel is reflected in research that has compared the relative validity of performance-based measures with self-report. In a meta-analysis of 61 studies in which the IAT and self-report indicators were used to predict the same or similar criteria, Poehlman et al. (2005) found a mean correlation for the IAT of .27 and for indicators of mental experience of .35. Poehlman et al. concluded that the IAT was most useful as an adjunct to self-description when the criteria were socially

sensitive or difficult to control consciously. Similar conclusions have been drawn for the Rorschach as well (e.g., Grossman, Wasyliv, Benn, & Gyoerkoe, 2002; Meyer & Archer, 2001).

Despite the similarities, the Rorschach can be distinguished from the IAT and related instruments in two important ways. The first is in terms of the trade-off between bandwidth and fidelity (Cronbach & Gleser, 1965), an issue that is not pertinent to the central argument of this article but I will discuss further in the Conclusion. The second has to do with the theoretical framework in which the two instruments function. Specifically, it is hypothesized that the effectiveness of the Rorschach and the Comprehensive System is undermined by the degree to which it is based on two problematic, and to some extent outmoded, beliefs. The first suggests that empirical methods can identify measures that are both subtle and valid. The second is a traditional psychoanalytic perspective on the nature of mental processes. Although the discussion of these issues and their implications focus primarily on variables from the Comprehensive System, the Rorschach Oral Dependency (ROD; Masling, Rabie, & Blondheim, 1967) scale proves useful for demonstrating several points, in part because it is one of the Rorschach indicators with the strongest empirical support (e.g., Bornstein, 1999).

#### INDIRECT MEASUREMENT

Many Rorschach scales have been developed on strictly empirical grounds without considering the degree to which the scale demonstrated a clear correspondence with the construct it ostensibly reflected. In fact, lack of correspondence has historically been considered a desirable feature for a Rorschach scale. Oftentimes the original empirical base for a Rorschach variable has even been informal, consisting of clinical assertions about an interesting covariation between test behavior and person attributes (Meyer, 1996).

Exner (2002) deserves acknowledgement for his recognition that all scores in the Comprehensive System should be based on formal empirical validation. However, the manner in which this goal was accomplished demonstrates excessive faith in the power of empirical approaches. Exner (2002) described numerous instances in which coding rules were modified to enhance interrater agreement, the interpretation of scores was derived directly from empirical studies, or composite variables such as the Depression Index were created based on a discriminant function analysis. In many instances, the results were never cross-validated or considered in light of variable–construct relationships. Critics of the Rorschach have raised concerns about the methodological rigor of Exner’s (e.g., 2002) research (e.g., Wood et al., 1996), but that is not the focus of this discussion. The goal here is to point out that unreplicated empirical results without consideration of the degree to which the scale is intuitively related to the construct is a flawed approach to scale development. The link between variable and construct has been discussed in terms of item prototypicality, trait indicativity, content validity, and subtle versus obvious items (Broughton, 1984; Cronbach & Meehl, 1955; Johnson, 2004; Meehl, 1945), but I generally refer to them here as having to do with indirect versus direct measurement.

Historically, interest in the use of measures that are indirectly related to the construct they are intended to indicate can be traced to the 1920s when psychologists first became

seriously concerned that people may not respond accurately to transparent items (Hartshorne & May, 1928). The popularity of the Rorschach in the United States in part originated out of these concerns. They also played a significant role in the development of the MMPI, for which items were chosen primarily on the basis of statistical analyses rather than conceptual grounds. For example, if there was a significant difference between depressives and members of the normative sample in responding to an item, that item became part of the MMPI depression scale keyed according to the response alternative more commonly chosen by those with depression. The result was a set of items to which those with depression responded differently than the normative sample regardless of any motivation among them to misrepresent themselves.

What is relevant to this discussion was the discovery that many of the items meeting the empirical criterion had no direct relevance to the target construct for the scale. Wiener (1948) hypothesized that these so-called subtle items should be relatively impervious to faking. Meehl (1945) further suggested that the subtle items could inform theory building by revealing unforeseen features of the construct.

This optimism about the potential for items with no apparent connection to the construct ultimately proved unfounded, in part because of a simplistic belief that the empirical effects were so robust as to obviate the need for cross-validation. Jackson's (1971) critique of atheoretical empirical scales proved to be a turning point in psychologists' faith in the naive empirical approach. Jackson (1971) proposed that

Personality measures will have broad import and substantial construct validity to the extent, and only to the extent, that they are derived from an explicitly formulated, theoretically based definition of a trait. (p. 232)

Although Jackson did not completely discount the possibility of creating good subtle items based on conceptual grounds, he hypothesized that most such items chosen empirically for inclusion in the MMPI scales in fact represented Type I errors and would have been eliminated with cross-validation.

Numerous studies have been conducted evaluating the criterion-related validity of direct versus indirect items (e.g., Campbell & Mehra, 1958; Goldberg & Slovic, 1967; Johnson, 2004; Norman, 1963; Osberg & Harrigan, 1999; Paunonen, 1984; Weed, Ben-Porath, & Butcher, 1990; Wrobel & Lachar, 1982), and they have consistently supported Jackson's (1971) hypothesis. In each case, relationships with relevant criteria consistently covary positively with ratings of item directness. What is particularly relevant to the issue of cross-validation is that several of these studies have found both direct and indirect stimuli significantly predicted a criterion in a derivation sample, but only the items rated as direct indicators survived cross-validation. Although they are theoretically more susceptible to faking than indirect items, direct items have consistently proven to be better predictors—and the only reliable predictors—of reasonable criteria. Empirically selected indirect items are unlikely to demonstrate reliable effect sizes when the analyses are cross-validated.

Indicators that clearly correspond with the constructs they are thought to reflect have another advantage in terms of the construct validation of an instrument. A core component of construct validation involves clarifying the nature of the construct

represented by an observed scale (Embretson [Whitely], 2003; McGrath, 2005). Cronbach and Meehl (1955) thought the pattern of relationships with other observed variables would inform psychologists about the matrix of latent relationships within which the construct exists, a matrix they referred to as the nomological net, and this would enhance understanding of the construct itself. This in turn would permit more refined measures of the construct, a process Cronbach and Meehl referred to as "bootstrapping." In fact, the use of empirical evidence to bootstrap one's knowledge of the construct is difficult if not impossible to achieve unless the indicator demonstrates a strong conceptual link with the construct of interest.

For example, the race version of the IAT pairs African American and Euro-American faces with positive versus negative words. Research has consistently demonstrated a tendency among Americans to associate Black faces more strongly with negative words, and this effect is only mildly related to self-reported racism (Cunningham, Preacher, & Banaji, 2001). Despite lack of convergence with self-report, the nature of the task created a strong presumption that the race IAT reveals something interesting about reactions to members of different racial groups. Because this link is so intuitive, researchers have been able to generate fairly sophisticated hypotheses demonstrating, for example, that the effect is not due to familiarity (Dasgupta, Greenwald, & Banaji, 2003) or to bias against the color black (Smith-McLallen, Johnson, Dovidio, & Pearson, 2006), but racial bias on the IAT does correlate with reasonable indicators of subtle forms of prejudice such as attributions about behavior and treatment choices by physicians (Lane, Banaji, Nosek, & Greenwald, 2007).

Nosek, Greenwald, and Banaji (2007) wrote that "the IAT has enjoyed a period of sustained empirical use during which its creators, developers, and users have remained relatively calm about the absence of an established cognitive model of performance at the task that generates the IAT measure" (p. 283).<sup>2</sup> Even without being able to articulate the construct fully, the intuitive quality of the IAT provides a conceptual foundation for studies that refine the understanding of the construct in important ways, so that a cumulative body of research about the meaning of the IAT is possible.

In contrast, Exner proposed an elevated score on his Egocentricity Index "suggests that the individual tends to be much more involved with himself or herself than are most others" (2000, p. 257). The Egocentricity Index is based on the number of Pair responses in which the same object is identified on both sides of the inkblot and the number of Reflection responses in which an object on one side of the inkblot is reflected on the other.<sup>3</sup> In fact, no intuitive rationale exists for associating Pair responses with the construct self-involvement. Without such an association, all sorts of questions central to understanding the construct validity of the indicator are unanswerable: why this variable predicts self-involvement, whether a different weighting of Pair responses in the index would offer a better approach to gauging self-involvement, under what circumstances Pair

<sup>2</sup>Similarly, items that are intuitively related to the constructs they are thought to reflect are used successfully in self-report measurement even though some questions about their use remain unresolved, such as the degree to which these items reflect self-presentation versus self-disclosure (Johnson, 1981).

<sup>3</sup>A glossary that includes all Rorschach coding and scoring terms used in this manuscript may be found in Anonymous (2007).

responses predict self-involvement (hypotheses about moderator effects), what aspects of self-involvement are best predicted by Pair responses, and so forth. The reliance on Rorschach variables of uncertain meaning compromises attempts at improving the construct validity of those variables (Meyer, 1996; Widiger & Schilling, 1980).

In the years prior to the ascendance of the Comprehensive System, Weiner (1966, 1977) has criticized the strict empirical approach to developing Rorschach measures and has highlighted the importance of a conceptual model for Rorschach scores and their interpretation. Unfortunately, Weiner's (1966, 1977) message was lost in the groundswell of support for the Comprehensive System. The evidence is clear: Indicators that are directly linked to constructs of interest are superior whether evaluated in terms of maximizing the prediction of important criteria or maximizing the understanding of the scale-construct relationship.

#### TRADITIONAL PSYCHOANALYTIC PERSPECTIVE ON MENTAL PROCESSES

Whereas some Rorschach variables can be traced to the strict empiricist tradition, others are rooted in psychoanalytic assumptions about the nature of mental activity. For example, Rorschach responses involving food have historically been interpreted as evidence of dependency issues (Exner, 2000; Masling et al., 1967), an interpretation that has its origins in the psychoanalytic hypothesis that dependency reflects fixation in the oral stage of development.

A predilection toward using psychoanalytic theory to understand respondent behavior on performance-based personality measures made sense when psychoanalytic theory offered the only cogent account of unconscious mental processes. In recent years, however, dual-process theories have emerged as an alternative (e.g., Bargh & Chartrand, 1999; Barrett, Tugade, & Engle, 2004; McClelland et al., 1989; Weinberger, 1992; Wilson, Lindsey, & Schooler, 2000; Zajonc, 1980) and have quickly come to dominate experimental research in topics as diverse as making attributions, attention, social perception, memory, emotion, and personality. The popularity of the IAT is one manifestation of the success of dual-process theories.

Dual-process theories are consistent with psychoanalytic theory in recognizing that both mental experiences, associated in the literature with adjectives such as *conscious*, *explicit*, *supraliminal*, or *deliberate*, and mental processes, referred to as *unconscious*, *implicit*, *subliminal*, or *automatic*, are essential to a complete theory of mental functioning. For this reason, discussions of dual-process theories sometimes acknowledge a debt to Freud concerning his popularization of the concept of the unconscious. However, the two approaches use the concept in very different ways (Kihlstrom, 1999; Westen, 1998). Freud focused on unconscious information and processes that existed primarily because of an active process of repression as a defense against anxiety. In contrast, dual-process theories attribute the existence of unconscious activity primarily to the limitations of consciousness. Sensory inputs must surpass certain minimum thresholds before they can be consciously experienced, consciousness is limited in terms of the amount of information that can be manipulated simultaneously, mental processes may remain outside of consciousness simply because they have not been self-observed, and automatic processes can be more

efficient than deliberate processes. Although these findings do not preclude the hypothesis that unconscious activity can reflect defensive operations, they suggest unconscious activity largely occurs for benign reasons and, as noted previously, can be more accessible to conscious evaluation than psychoanalytic theory would traditionally suggest.

The psychoanalytic tradition continues to influence Rorschach practice in at least two ways. The more obvious influence is evident in interpretive strategies that reflect uncorroborated psychoanalytic hypotheses. The Food response is a good example of this dubious practice because even many psychodynamically oriented psychologists question the validity of the psychosexual model. Fortunately, the Comprehensive System offers few examples of variables that depend on traditional psychoanalytic thinking for their intuitive justification.

The more subtle, but also more pervasive, influence derives from the psychoanalytic assumption that mental activity is strongly influenced by conflicts and wishes, and verbal behavior can largely be interpreted as symbolic representations of those issues. This assumption remains evident in a tendency to interpret responses in a manner that is superficially direct, and appealing for that reason, but does not provide the most straightforward explanation for the response.

For example, it was noted earlier that Exner (2000) associated even one Reflection response with self-focusing, a link that would seem to make intuitive sense from a literary or psychoanalytic perspective. Whether it makes intuitive sense from the perspective of typical mental processes is another matter. This interpretation assumes that self-focused individuals demonstrate an automatic focus on acts of physical self-observation, an untested assumption. A more parsimonious explanation would suggest the Reflection response is a means of resolving the perceptual challenge created by the symmetry of the blot (see Horn, Meyer, & Mihura, in press). The more intriguing hypothesis that it indicates self-focus may only be justified if there are additional signals of preoccupation. These might include multiple Reflection responses (although even in this case, a more reasonable explanation might be a preoccupation with resolving the symmetrical character of the blot, for which the Reflection strikes the respondent as a particularly clever strategy) or embellishment of the response in such a way as to suggest a preoccupation with self-observation (although in some cases, this embellishment might suggest the Reflection is more accurately interpreted as a concern about how the respondent is perceived by others). Using the terminology from the previous section, psychoanalytic theory sometimes can lead the test user to perceive a direct link between a test behavior and a construct and ignore a more direct—but perhaps more mundane—explanation.

#### IMPLICATIONS FOR THE RORSCHACH

##### *Core Hypotheses*

The preceding discussion suggests two hypotheses concerning the Rorschach that provide a framework for subsequent discussion. First, Rorschach variables with a direct intuitive link to the construct they are thought to represent will probably prove to be the most valid and will contribute the most to a cumulative research approach to the instrument. The construct need not be fully specified, but the relationship should be clear. Second, the psychoanalytic approach to the interpretation of symbols has resulted in interpretive statements that ignore more direct

and parsimonious explanations of test behavior. An important auxiliary principle derived from these two hypotheses suggests that response elements that are structurally equivalent in terms of current Comprehensive System coding may in fact reflect different underlying mental processes, and contextual elements will play an important role in hypotheses about what processes in fact are operating. In the remainder of this article, I focus on some implications of these hypotheses for Rorschach practice. The presentation is more conjectural than has been the case up to this point, with the goal of initiating discussion about the use of the Rorschach rather than drawing final conclusions.

### *Cognitive Processes, Associative Processes, and Extra-Test Behavior*

To understand how an intuitive link can be hypothesized between a Rorschach variable and a construct, it is important to consider what mental processes might underlie a response. Literature on the Rorschach has often distinguished between three classes of processes (Weiner, 1977, 1994). The first can be broadly referred to as cognitive. These include the perceptual processes used to resolve the problem of “what might this be?” and the logical and linguistic processes used to formulate the statement of the percept. Good examples of intuitively linked indicators of such processes from the Comprehensive System include Form Quality as an indicator of deviant perceptual tendencies and Deviant Verbalizations as evidence of difficulties in the verbal expression of ideas.

Other activity seems to reflect associational processes, which is the aspect of Rorschach test behavior most consistent with its traditional labeling as a projective instrument. The discussion of Reflection responses in the previous section suggested a core problem in the identification of associational activity may be the discrimination between the perceptual processes used to identify a percept in the inkblot and an associational output (see also Rosegrant, 1984). Several heuristics can be suggested to exemplify this process, although exceptions can be noted to each. One is that an associational process will tend produce a recurring theme and/or greater detail or elaboration. A second is that nouns (content) are perceptual, whereas adjectives and adverbs (descriptions or qualifiers) are associational.

An example of a potentially interesting associational variable is reflected in the use of Texture (the use of Shading to suggest a tactile impression) to indicate issues surrounding interpersonal closeness, although I evaluate this variable in greater detail following. The attribution of Texture is explicitly descriptive. Even the rare pure Texture response is intended to be descriptive, although it is descriptive of the blot itself rather than some object the blot is thought to resemble. A second example of an associational variable is the Morbid response (indications the percept is dead, damaged, or associated with dysphoric affect). This is in contrast with Form Quality, which addresses the fit of the object to the blot location and is clearly more of a perceptual processing issue (although disturbing associations hypothetically can lead to the addition of distorted features to the percept).

Consideration of the distinction between cognitive and associational factors in a response generates two intriguing hypotheses about Rorschach administration and interpretation. First, the standard Comprehensive System inquiry question “what makes it look like that?” assumes the central mental process underlying

the response is perceptual. If there is reason to believe a particular element of a response is more interesting because of an associational process, however, it might be more productive to add questions of the type “tell me more about it being \_\_\_” or “what do you think of when you think of \_\_\_?” (see Aronow, Reznikoff, & Moreland, 1994).

A second hypothesis derived from considering the role of both cognitive and associational processes would suggest that rather than focusing on specific contents or determinants, the interpretation of associations may benefit more from an analysis of the language used by the respondent. This approach potentially permits the coding of personal attributes never formalized within the confines of the Comprehensive System such as the use of demeaning or dramatic descriptors.

The third potentially interesting component of responding is extra-test behaviors, including speech patterns. Of course, extra-test behavior is potentially interesting no matter what the method of assessment. The person who starts writing notes in the margin of the MMPI answer sheet is potentially revealing something more important than the information gleaned from the test itself. However, the interactive quality of the Rorschach administration combined with the ambiguous, vaguely evocative, and anxiety-provoking qualities of the stimulus objects enhance the potential for interesting extra-test behaviors. The Comprehensive System has only a few scores in which respondent behavior is important, such as Deviant Responses (odd or circumstantial phrasing) and the use of rubbing the card in relation to coding a Texture response (Exner, 2000), but the issue has been discussed more extensively in other literature on the method. For example, Weiner (2003) listed five categories of Rorschach extra-test behavior, including unusual card handling, comments, personalized responses, expressive style, and interpersonal style. Clinical use of the Rorschach suggests extra-test behaviors can offer conceptually intuitive indicators of interpersonal style and may merit greater inclusion in formal coding. Ideally, a comprehensive system for the Rorschach would encompass all aspects of behavior during the test administration.

### *Considering Context*

The core problem when using the Rorschach is that the respondent's behavior is not governed by the simplifying assumptions of the psychologist. What looks like the same type of response can potentially reflect a variety of mental processes. For example, multiple Personalized Answers (answers justified or clarified by reference to personal knowledge) can result because the respondent is uncomfortable with the Rorschach or with testing in general, is insecure about their abilities, is attempting to assert superiority over the tester through expression of their personal knowledge, or is demonstrating narcissistic tendencies (Weiner, 2003, p. 228). Greater consideration of context is similarly useful for understanding some of the cognitive special scores, which can reflect immaturity, subcultural factors, education, and various other factors. The incorporation of alternative hypotheses for interpreting a certain type of response, and factors that can be used to distinguish between those alternatives, into the interpretive algorithms for the Comprehensive System could dramatically enhance the validity of the Rorschach.

Exner's (2002) coding rules at times complicate the process of considering context. The coding rules were influenced by multiple factors including the historical roots of the code and

concern for interrater agreement. In many cases, these considerations trumped establishing an intuitive link between the signifier and what is signified.

Texture responses offer a good example of this phenomenon. Exner (2000) interpreted Texture based on Shading as having to do with “needs for closeness and openness to close emotional relations,” (p. 314) and multiple Texture responses are interpreted as evidence of unfulfilled needs for closeness. Texture is an example of a code where both conceptual considerations and empirical evidence support the hypothesized interpretation (Marsh & Viglione, 1992). However, the intuitive interpretation only makes sense as the result of an associative process when a Texture response reflects comforting attributes such as warmth, softness, or fuzziness. For reasons of maximizing reliability, Texture is instead scored any time there is a tactile impression, including wetness, hardness, or roughness. References to hardness or roughness hypothetically could indicate the denial of a need for closeness, but it is intuitively unlikely that they ever reflect openness to being close, and there is no intuitive basis for including wetness in the list. One might even hypothesize that references to hardness at times suggest an unconflicted tendency toward distancing. Alternatively, any Texture response could suggest a perceptual process. Again, the correct interpretation probably depends on the frequency and elaboration of Texture responses. Furthermore, a reference to warm fur based on the form demand of the blot is potentially more indicative of the relevant construct than the perception of wetness based on shading, depending on how the respondent qualifies the response. The restriction of the code to shading-based Texture responses can be traced directly to Klopfer’s (1938) original comments on the Shading response rather than a conceptual analysis of the link between the code and the construct.

A second example involves the Morbid response. Consider two similar responses to Card 1 that would warrant coding as Morbid:

1. This is a butterfly. Its wings are ripped and tattered, and it doesn’t have very long to live.
2. This is a butterfly. I don’t know what to make of these white spaces, I don’t know any kind of butterfly with white spots on its wings quite like that. They really shouldn’t be there, but I guess its wings are ripped.

The first is intuitively indicative of the negativity that is associated with Morbid responses, as reflected in the reference to death and the redundancy of *ripped* and *tattered*. The second seems more indicative of an intellectualized or obsessional quality in responding to the blot as reflected by the equivocation about the explanation. What is important is not the reference to damage but the manner in which the respondent qualifies that reference. It is an open question whether it is possible to refine the coding rule for Morbid responses in such a way as to allow discrimination between these two examples without compromising interrater reliability.

One of the important contributions the Comprehensive System has made to the Rorschach is standardization (e.g., Meyer, Mihura, & Smith, 2005), and it is important to clarify that the point here is not to undermine consistency in the way the instrument is administered, scored, or interpreted. However, a conceptual analysis of Rorschach behavior suggests it is far more complex than the existing interpretive guides would suggest.

Superficial similarities in responses can indicate very different mental processes. A maximally valid interpretive approach to the Rorschach would need to address that complexity explicitly.

#### *Codes With Questionable Content Validity*

There are several Comprehensive System variables for which there would seem to be either no credible conceptual model that explains the hypothesized indicator-construct link, or the link is so distant that one should question how frequently it occurs in practice. The following admittedly subjective list offers some examples:

- Diffuse Shading (use of Shading to indicate variation) and Vista (use of Shading to indicate depth) responses.
- Animal and Inanimate Movement responses.
- Pairs and Reflections.
- Form Dimensionality (use of Form to indicate depth).
- Food responses.

Prior research has tended to support raising concerns about the validity of some of these variables (e.g., Eells & Boswell, 1994; Nezworski & Wood, 1995; Viglione & Exner, 1983), and others remain almost untouched in the published literature. In the previous section, I did suggest the possibility that a Vista or Diffuse Shading response can at times reflect some important respondent attribute but only when the language used in connection with the response is consistent with the attribute. For example, a Vista response accompanied by a reference to a deep, dark hole would hypothetically be more predictive of depression or suicidal intent than a reference to the Grand Canyon observed from above based on both Form and Shading. To state this proposition another way, the focus on determinant or content codes is potentially misleading to the extent it deemphasizes the specific manner in which the determinant or content is used by the respondent. Notice that what has been suggested here is not that some Comprehensive System variables such as Texture are valid, whereas others such as Vista are not; rather, the hypothesis is that Texture responses are more consistently reflective of the proposed latent construct than is true of Vista responses.

#### *Excessive Aggregation*

Many of the key variables in the Comprehensive System are composites of other variables. Examples include the Suicide Constellation, Depression Index, Perceptual-Thinking Index, D Score, and Coping Deficit Index. Exner (2002) selected the components of these indexes on the basis of (often unreplicated) empirical analyses. As McGrath (2005) noted, aggregated scores are useful because their reliability tends to exceed that of its elements, and their criterion-related validity tends to exceed the mean validity of those elements. This justification for generating indexes makes particular sense in the case of scales such as those from the MMPI in which the elements are individual items. It is not as compelling in the case of the Comprehensive System because the elements are already based on aggregates such as the total number of Passive Movement or Space (use of white space) responses.<sup>4</sup>

<sup>4</sup>As a side note, the conversion of the elements to dichotomous outcomes before their aggregation rather than standardization and summation could also compromise the validity of the composite.



In cases in which the composite includes some elements that demonstrate relatively poor criterion-related validity, it is reasonable to expect that a subset of the elements can outperform the aggregate. If some or most of the elements do not demonstrate an intuitive link to the construct associated with the index and so are unlikely to cross-validate, the index as a whole may prove to be a valid indicator of some construct but less valid than a composite based only on the most valid elements.

The indexes vary in the degree to which they are comprised of intuitively linked elements. The Perceptual-Thinking Index is probably the best in this regard. In contrast, the Suicide Constellation consists of a very diverse set of variables that differ markedly in terms of their intuitive link to the construct. In fact, it may be hypothesized that a tally of the number of responses in which the language suggests negative affect could prove more valid than the total index. This argument also suggests the possible value of decomposition research for Comprehensive System index scores.

Such studies are rare. One recent example is available for the ROD, which is a composite of variables thought to be predictive of dependent tendencies. Derived from psychoanalytic theory, the scale involves counting the frequency of both oral references (references to food and the mouth) and of imagery that implies a dependent relationship. Bornstein (2007a) described two studies in which he had examined the impact of behavioral manipulations on performance on the ROD. In both cases, ROD total scores varied as expected. However, when he examined dependency and oral imagery separately, the effect was only significant for the intuitively relevant dependency imagery. The results highlight the potential for more valid composites that focus on more intuitively meaningful components of Rorschach indexes, especially if coding rules for those components are enhanced in light of the construct-code relationship.

### *The Assumption of Bipolar Structure*

The proposition that superficially similar Rorschach variables can reflect several mental processes leads to an important correlate, which is that the meaning of a Rorschach variable need not be consistent across the entire range of outcomes. For example, a protocol with fewer than the typical number of Popular (commonly occurring) responses is interpreted as evidence of an unconventional approach to interpreting stimuli, whereas a high rate of Popular responses is interpreted as evidence of a person "overly involved with the detection of cues related to socially expected or acceptable behaviors" (Exner, 2000, pp. 184–185). The former interpretation seems more intuitively appealing than the latter. There are in fact at least three possible models for understanding the number of Popular response (see also Weiner, 2003, p. 111):

1. The variable may be bipolar unidimensional. That is, the two extremes suggest opposite conclusions about the respondent. This is the option implied by interpreting a high number of Populars as evidence of excessive conventionality.
2. The variable may be unipolar unidimensional. This would be the case if a low number of Populars suggests unconventionality, but a high number has no interesting interpretation.
3. The variable may be unipolar bidimensional. This would be the case if a low number of Populars suggests unconvention-

ality, whereas a high number suggests something distinct from conventionality.

This analysis suggests the importance of considering the possible intuitive explanations for performance at each extreme of Rorschach variables.

### CONCLUSIONS

Although the IAT is not without its critics (e.g., Blanton, Jaccard, Gonzales, & Christie, 2006), the degree of interest it has garnered suggests there is still a good future for the performance-based assessment of personality and psychopathology. The purpose of this article was to point a direction for future work involving the Rorschach by comparing it with the parallel class of instruments that has emerged out of dual-process theories, of which the IAT has served as the exemplar. Current Rorschach practice, including but not limited to the Comprehensive System, still reflects the influence of both excessive faith in empiricism and unverified psychoanalytic assumptions. An approach emphasizing variables that are intuitively meaningful, and informed by reasonable hypotheses about the mental activities underlying those variables, offers an alternative with potentially important implications for Rorschach research and clinical use.

This discussion suggests that even after 80 years, it is still too early to evaluate the true potential of the Rorschach as an assessment instrument. This is not the fault of the instrument itself. The Rorschach method of data acquisition has proven to be a projective stimulus in its own right, used in ways reflecting the presuppositions and postulates of its times (a statement that to a lesser extent can be applied to the MMPI as well: cf. Meehl, 1945, to Weed et al., 1990). Fortunately, psychologists have become much more knowledgeable about how to design valid indicators and how to study automatic processes. That knowledge should guide future development of the Rorschach.

Even if the recommendations outlined in this article are implemented, it is worth noting that the Rorschach will continue to demonstrate certain limitations that should be considered in its use. First, it was mentioned earlier that the Rorschach is a broadband indicator of personality, and its continued popularity in part stems from this attribute. The CPT and TOVA were designed specifically to tap into a very small set of constructs relevant to personality and psychopathology. The IAT similarly addresses a single pair of associations at a time, such as alcohol positive and alcohol negative. In contrast, the ambiguous nature of the Rorschach stimuli makes it possible for a much broader set of personal tendencies to emerge during the testing. It is the possibility of such breadth that makes the Rorschach interesting to practitioners of clinical assessment in a way that the CPT or IAT could never be. In particular, if the cost of administering an instrument is computed in proportion to the number of variables into which it potentially offers insight, the clinician may be justified in concluding the cost of administering the Rorschach is justified. This can be particularly true when an assessment is conducted with exploration rather than confirmation in mind, for purposes of description rather than prediction.

At the same time, its fidelity for certain constructs will never match that of more narrow-band instruments. For example, even if a high frequency or a dramatic elaboration of Texture responses proves a valid indicator of issues with intimacy, the proposition that an individual with intimacy issues is driven to focus on the textural potentialities of the blot reflects unverified

psychoanalytic assumptions. In contrast, an IAT that uses terms reflecting intimacy versus aloofness with self-referential versus neutral terms directly confronts the respondent with the issue of self in relation to concepts reflecting interpersonal distance. Although the ambiguity of the Rorschach permits a broader set of stylistic elements to emerge than the IAT, the same ambiguity means the pull for any one of these elements is not sufficiently strong to assure a high likelihood of its emergence. It is important to note that the basic instructions "what might this be?" suggests the Rorschach's inherent pull for perceptual processes is likely to be stronger than that for associational or extra-test behaviors. As noted earlier, this disparity may be at least partially ameliorated by greater use of prompts during the inquiry that pull for associational processes.

A second factor compromising the Rorschach's potential for fidelity has to do with the assumption that stylistic responding to the inkblots reflects the respondent's general approach to environmental, and particularly interpersonal, stimuli. This assumption is consistent both with traditional psychoanalytic thinking, which assumes the person is driven to express core conflicts and defenses across stimulus situations, and some early literature touting empirically derived measurements. Berg (1957) presented the most explicit formulation with regard to the latter. Berg proposed that stylistic differences between normal and deviant populations are likely to emerge in any stimulus situation, and so valid discriminators can be developed through strict empirical selection from any class or modality of stimuli. Berg's speculation was subsequently refuted by several of the studies cited earlier on direct versus indirect measurement (Goldberg & Slovic, 1967; Norman, 1963).

More recent literature on performance-based assessment suggests quite the opposite: Deviant test behavior is often specific to the class of stimuli used. CPT scores based on visual stimuli do not correlate very well with attentional indicators using other sense modalities (e.g., Borgaro et al., 2003). Similarly, Nosek et al. (2007) reviewed studies comparing the IAT with other types of indicators of automatic processes and found correlations were quite poor, and in some cases negative. Nosek et al. concluded the results were at least in part due to differences in the cognitive processes underlying responses to different types of stimuli. Although existing validation evidence demonstrates that respondent behavior to the Rorschach generalizes to other stimulus situations, the extent of stimulus specificity in Rorschach responding has simply not been adequately considered. The limited consideration of stimulus specificity in the Rorschach literature undoubtedly reflects the traditional emphasis on psychoanalytic rather than dual-process explanations for Rorschach behavior.

It is important to note that this article is not intended to imply that a clear conceptual correspondence between a Rorschach variable and a construct is sufficient to ensure the validity of the variable. As I suggested in the section on psychoanalytic theory, variable interpretations that seem intuitive from a certain theoretical perspective are not necessarily so. Similarly, the prior discussion has suggested that intuitive interpretations are compromised to the extent that Rorschach responding is stimulus specific. The existence of a conceptual link does not abrogate the need for validity research, which is why research demonstrating scores on the IAT predict important outcomes remains essential despite the intuitive attractiveness of the instrument (Lane et al., 2007). Unfortunately, much of the research

on the Comprehensive System is observational in nature, comparing preexisting groups that potentially differ in a variety of ways. One validity paradigm that has proven particularly useful for other performance-based measures (e.g., Atkinson & McClelland, 1948; Rudman, Dohn, & Fairchild, 2007) as well as the Rorschach (e.g., Bornstein, 2007a) compares individuals responding after alternative experimental manipulations as a means of minimizing differences between groups.

This article has only barely touched on other instruments that rely on ambiguous stimuli. However, many of the recommendations offered here are equally applicable to storytelling techniques, figures drawings, and sentence completion tests (although the latter represent a hybrid between performance-based and self-descriptive measurement strategies). It has intentionally raised more questions than it has answered, in that the goal was to initiate a discussion about the future direction for the Rorschach. Despite Exner's (2002) groundbreaking efforts in terms of putting the instrument on a firmer scientific basis, the Rorschach is still embedded within a set of presuppositions for which research has suggested better alternatives. Applying science on optimal scale development and on the measurement of mental processes, and lessons learned about performance-based personality assessment using instruments such as the IAT, offers the potential for improving Rorschach science and practice and for reinvigorating the image of this most intriguing instrument within scientific circles.

#### ACKNOWLEDGMENTS

I am deeply grateful for the advice and encouragement provided by Stephen Armeli, Robert Bornstein, Robert Erard, Howard Garb, Donald Viglione, Irving Weiner, and an anonymous reviewer. The challenges their feedback presented to me dramatically improved the quality of this article. I hope they find the final result worthy of their efforts. Portions of this article were presented at the 2004 Midwinter Meeting of the Society for Personality Assessment, Miami, FL.

#### REFERENCES

- Anonymous. (2007). Glossary of Comprehensive System scores. *Journal of Personality Assessment, 89*, S217-S220.
- Archer, R. P., & Krishnamurthy, R. (1993a). Combining the Rorschach and the MMPI in the assessment of adolescents. *Journal of Personality Assessment, 60*, 132-140.
- Archer, R. P., & Krishnamurthy, R. (1993b). A review of MMPI and Rorschach interrelationships in adult samples. *Journal of Personality Assessment, 61*, 277-293.
- Archer, R. P., & Newsom, C. R. (2000). Psychological test usage with adolescent clients: Survey update. *Assessment, 7*, 227-235.
- Aronow, E., Reznikoff, M., & Moreland, K. (1994). *The Rorschach technique: Perceptual basis, content interpretation, and applications*. Boston: Allyn & Bacon.
- Atkinson, J. W., & McClelland, D. C. (1948). The projective expression of needs: II. The effect of different intensities of the hunger drive on thematic apperception. *Journal of Experimental Psychology, 38*, 643-658.
- Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist, 54*, 462-479.
- Barrett, L. F., Tugade, M. M., & Engle, R. W. (2004). Individual differences in working memory capacity and dual-process theories of the mind. *Psychological Bulletin, 130*, 553-573.
- Berg, I. A. (1957). Deviant responses and deviant people: The formulation of the deviation hypothesis. *Journal of Counseling Psychology, 4*, 154-161.

- Blanton, H., Jaccard, J., Gonzales, P. M., & Christie, C. (2006). Decoding the Implicit Association Test: Implications for criterion prediction. *Journal of Experimental Social Psychology, 42*, 192–212.
- Block, J. (2002). *Personality as an affect-processing system: Toward an integrative theory*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Borgaro, S., Pogge, D. L., DeLuca, V. A., Bilginer, L., Stokes, J., & Harvey, P. D. (2003). Convergence of different versions of the Continuous Performance Test: Clinical and scientific implications. *Journal of Clinical and Experimental Neuropsychology, 25*, 283–292.
- Bornstein, R. F. (1999). Criterion validity of objective and projective dependency tests: A meta-analytic assessment of behavioral prediction. *Psychological Assessment, 11*, 48–57.
- Bornstein, R. F. (2002). A process dissociation approach to objective-projective test score interrelationships. *Journal of Personality Assessment, 78*, 47–68.
- Bornstein, R. F. (2007a). Might the Rorschach be a projective test after all? Social projection of an undesired trait alters Rorschach Oral Dependency Scores. *Journal of Personality Assessment, 88*, 354–367.
- Bornstein, R. F. (2007b). Toward a process-based framework for classifying personality tests: Comment on Meyer and Kurtz (2006). *Journal of Personality Assessment, 89*, 202–207.
- Broughton, R. (1984). A prototype strategy for construction of personality scales. *Journal of Personality and Social Psychology, 47*, 1334–1346.
- 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.
- Camara, W. J., Nathan, J. S., & Puente, A. E. (2000). Psychological test usage: Implications in professional psychology. *Professional Psychology: Research and Practice, 31*, 141–154.
- Campbell, D. T., & Mehra, K. (1958). Individual differences in evaluations of group discussions as a projective measure of attitudes toward leadership. *Journal of Social Psychology, 47*, 101–106.
- Conners, C. K. (1994). *The Conners Continuous Performance Test*. Toronto, Ontario, Canada: Multi-Health Systems.
- Cronbach, L. J., & Gleser, G. C. (1965). *Psychological tests and personnel decisions*. Urbana: University of Illinois Press.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*, 281–302.
- Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). Implicit attitude measures: Consistency, stability, and convergent validity. *Psychological Science, 12*, 163–170.
- Dasgupta, N., Greenwald, A. G., & Banaji, M. R. (2003). The first ontological challenge to the IAT: Attitude or mere familiarity? *Psychological Inquiry, 14*, 238–243.
- De Houwer, J., Crombez, G., Koster, E. H. W., & De Beul, N. (2004). Implicit alcohol-related cognitions in a clinical sample of heavy drinkers. *Journal of Behavior Therapy and Experimental Psychiatry, 35*, 275–286.
- De Houwer, J., & Moors, A. (2007). How to define and examine the implicitness of implicit measures. In B. Wittenbrink & N. Schwarz (Eds.), *Implicit measures of attitudes* (pp. 179–194). New York: Guilford.
- Dickersin, K. (2005). Publication bias: Recognizing the problem, understanding its origins and scope, and preventing harm. In H. R. Rothstein, A. J. Sutton, & M. Borenstein (Eds.), *Publication bias in meta-analysis: Prevention, assessment and adjustments* (pp. 11–34). Chichester, England: Wiley.
- Eells, G. T., & Boswell, D. L. (1994). Validity of Rorschach inanimate movement and diffuse shading responses as measures of frustration and anxiety. *Perceptual and Motor Skills, 78*, 1299–1302.
- Embretson (Whitely), S. (1983). Construct validity: Construct representation versus nomothetic span. *Psychological Bulletin, 93*, 179–197.
- Exner, J. E., Jr. (1989). Searching for projection in the Rorschach. *Journal of Personality Assessment, 53*, 520–536.
- Exner, J. E., Jr. (2000). *A Rorschach workbook for the Comprehensive System*. Asheville, NC: Rorschach Workshops.
- Exner, J. E., Jr. (2002). *The Rorschach: A comprehensive system: Vol. 1. Basic foundations and principles of interpretation* (4th ed.). New York: Wiley.
- Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology, 54*, 297–327.
- Garb, H. N., Wood, J. W., Lilienfeld, S. O., & Nezworski, M. T. (2005). Roots of the Rorschach controversy. *Clinical Psychology Review, 25*, 97–118.
- Gawronski, B., Hofmann, W., & Wilbur, C. J. (2006). Are “implicit” attitudes unconscious? *Consciousness and Cognition: An International Journal, 15*, 485–499.
- Gawronski, B., LeBel, E. P., & Peters, K. R. (2007). What do implicit measures tell us? Scrutinizing the validity of three common assumptions. *Perspectives on Psychological Science, 2*, 181–193.
- Goldberg, L. R., & Slovic, P. (1967). Importance of test item content: An analysis of a corollary of the deviation hypothesis. *Journal of Counseling Psychology, 14*, 462–472.
- Greenberg, L. M. (1990). *Test of Variables of Attention (TOVA)*. Los Alamitos, CA: Universal Attention Disorders.
- Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102*, 4–27.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology, 74*, 1464–1480.
- Grossman, L. S., Wasylw, O. E., Benn, A. F., & Gyoerko, K. L. (2002). Can sex offenders who minimize on the MMPI conceal psychopathology on the Rorschach? *Journal of Personality Assessment, 78*, 484–501.
- Hartshorne, H., & May, M. A. (1928). *Studies in the nature of character: I. Studies in deceit*. New York: Macmillan.
- Hiller, J. B., Rosenthal, R., Bornstein, R. F., Berry, D. T. R., & Brunell-Neuleib, S. (1999). A comparative meta-analysis of Rorschach and MMPI validity. *Psychological Assessment, 11*, 278–296.
- Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the Implicit Association Test and explicit self-report measures. *Personality and Social Psychology Bulletin, 31*, 1369–1385.
- Horn, S. L., Meyer, G. J., & Mihura, J. L. (in press). Impact of card rotation on the frequency of Rorschach reflection responses. *Journal of Personality Assessment*.
- Hunsley, J., & Bailey, J. M. (1999). The clinical utility of the Rorschach: Unfulfilled promises and an uncertain future. *Psychological Assessment, 11*, 266–277.
- Jackson, D. N. (1971). The dynamics of structured personality tests: 1971. *Psychological Review, 78*, 229–248.
- Johnson, J. A. (1981). The “self-disclosure” and “self-presentation” views of item response dynamics and personality scale validity. *Journal of Personality and Social Psychology, 40*, 761–769.
- Johnson, J. A. (2004). The impact of item characteristics on item and scale validity. *Multivariate Behavioral Research, 39*, 273–302.
- Kihlstrom, J. F. (1999). The psychological unconscious. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed.; pp. 424–442). New York: Guilford.
- Klopfer, B. (1938). The shading responses. *Rorschach Research Exchange, 2*, 76–79.
- Kubiszyn, T. W., Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland, K. L., et al. (2000). Empirical support for psychological assessment in clinical health care settings. *Professional Psychology: Research and Practice, 31*, 119–130.
- Lane, K. A., Banaji, M. R., Nosek, B. A., & Greenwald, A. G. (2007). Understanding and using the Implicit Association Test: IV. What we know (so far) about the method. In B. Wittenbrink & N. Schwarz (Eds.), *Implicit measures of attitudes* (pp. 59–102). New York: Guilford.
- Langer, F. (2004). Pairs, reflections, and the Ego: I. Exploration of a perceptual hypothesis. *Journal of Personality Assessment, 82*, 114–126.
- Lilienfeld, S. O., Wood, J. M., & Garb, H. N. (2000). The scientific status of projective techniques. *Psychological Science in the Public Interest, 1*, 27–66.
- Marsh, A., & Viglione, D. J. (1992). A conceptual validation study of the texture response on the Rorschach. *Journal of Personality Assessment, 58*, 571–579.

- Masling, J., Rabie, L., & Blondheim, S. H. (1967). Obesity, level of aspiration, and Rorschach and TAT measures of oral dependence. *Journal of Consulting Psychology, 31*, 233–239.
- Mayer, J. D. (2005). A tale of two visions: Can a new view of personality help integrate psychology? *American Psychologist, 60*, 294–307.
- McClelland, D. C., Koestner, R., & Weinberger, J. (1989). How do self-attributed and implicit motives differ? *Psychological Review, 96*, 690–702.
- McGrath, R. E. (2005). Conceptual complexity and construct validity. *Journal of Personality Assessment, 85*, 112–124.
- Meehl, P. E. (1945). The dynamics of “structured” personality tests. *Journal of Clinical Psychology, 1*, 296–303.
- Meyer, G. J. (1996). Construct validation of scales derived from the Rorschach Method: A review of issues and introduction to the Rorschach Rating Scale. *Journal of Personality Assessment, 67*, 598–628.
- Meyer, G. J., & Archer, R. P. (2001). The hard science of Rorschach research: What do we know and where do we go? *Psychological Assessment, 13*, 486–502.
- 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., Mihura, J. L., & Smith, B. L. (2005). The interclinician reliability of Rorschach interpretation in four data sets. *Journal of Personality Assessment, 84*, 296–314.
- Meyer, G. J., Riethmiller, R. J., Brooks, R. D., Benoit, W. A., & Handler, L. (2000). A replication of Rorschach and MMPI–2 convergent validity. *Journal of Personality Assessment, 74*, 175–215.
- Millon, T., Millon, C. M., Meagher, S., Grossman, S., & Ramnath, R. (2004). *Personality disorders in modern life* (2nd ed.). New York: Wiley.
- Nezworski, M. T., & Wood, J. M. (1995). Narcissism in the Comprehensive System for the Rorschach. *Clinical Psychology: Science and Practice, 2*, 179–199.
- Nock, M. K., & Banaji, M. R. (2007). Prediction of suicide ideation and attempts among adolescents using a brief performance-based test. *Journal of Consulting and Clinical Psychology, 75*, 707–715.
- Norman, W. T. (1963). Relative importance of test item content. *Journal of Consulting Psychology, 27*, 166–174.
- Nosek, B. A. (2005). Moderators of the relationship between implicit and explicit evaluation. *Journal of Experimental Psychology: General, 134*, 565–584.
- Nosek, B. A. (2007). Implicit-explicit relations. *Current Directions in Psychological Science, 16*, 65–69.
- Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2007). The Implicit Association Test at age 7: A methodological and conceptual review. In J. A. Bargh (Ed.), *Automatic processes in social thinking and behavior* (pp. 265–292). New York: Psychology Press.
- Olson, M. A., Fazio, R. H., & Hermann, A. D. (2007). Reporting tendencies underlie discrepancies between implicit and explicit measures of self-esteem. *Psychological Science, 18*, 287–291.
- Osberg, T. M., & Harrigan, P. (1999). Comparative validity of the MMPI–2 Wiener-Harmon subtle-obvious scales in male prison inmates. *Journal of Personality Assessment, 72*, 36–48.
- Parker, K. C. H., Hanson, R. K., & Hunsley, J. (1988). MMPI, Rorschach, and WAIS: A meta-analytic comparison of reliability, stability, and validity. *Psychological Bulletin, 103*, 367–373.
- Paunonen, S. V. (1984). Optimizing the validity of personality assessments: The importance of aggregation and item content. *Journal of Research in Personality, 18*, 411–431.
- Poehlman, T. A., Uhlmann, E. L., Greenwald, A. G., & Banaji, M. R. (2005). *Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity*. Unpublished manuscript, Department of Psychology, Yale University, New Haven, CT.
- Roediger, H. L. (1990). Implicit memory: Retention without remembering. *American Psychologist, 45*, 1043–1056.
- Rosegrant, J. (1984). Rorschach object relations and fantasy themes incorrectly scored as determinants. *Journal of Personality Assessment, 48*, 467–475.
- Rudman, L. A., Dohn, M. C., & Fairchild, K. (2007). Implicit self-esteem compensation: Automatic threat defense. *Journal of Personality and Social Psychology, 93*, 798–813.
- Schmidt, F. L. (1996). Statistical significance testing and cumulative knowledge in psychology: Implications for training of researchers. *Psychological Methods, 1*, 115–129.
- Schmukle, S. C., & Egloff, B. (2005). A latent state-trait analysis of implicit and explicit personality measures. *European Journal of Psychological Assessment, 21*, 100–107.
- Schultheiss, O. C. (2007). A memory-systems approach to the classification of personality tests: Comment on Meyer and Kurtz (2006). *Journal of Personality Assessment, 89*, 197–201.
- Smith-McLallen, A., Johnson, B. T., Dovidio, J. F., & Pearson, A. R. (2006). Black and white: The role of color bias in implicit race bias. *Social Cognition, 24*, 46–73.
- Vigilione, D. J., & Exner, J. E., Jr. (1983). The effects of state anxiety and limited social-evaluative stress on the Rorschach. *Journal of Personality Assessment, 47*, 150–154.
- Vigilione, D. J., & Hilsenroth, M. J. (2001). The Rorschach: Facts, fictions, and future. *Psychological Assessment, 13*, 452–471.
- Weed, N. C., Ben-Porath, Y. S., & Butcher, J. N. (1990). Failure of Wiener and Harmon Minnesota Multiphasic Personality Inventory (MMPI) subtle scales as personality descriptors and as validity indicators. *Psychological Assessment, 2*, 281–285.
- Weinberger, J. (1992). Validating and demystifying subliminal psychodynamic activation. In R. F. Bornstein & T. S. Pittman (Eds.), *Perception without awareness* (pp. 170–190). New York: Guilford.
- Weiner, I. B. (1966). *Psychodiagnosis in schizophrenia*. New York: Wiley.
- Weiner, I. B. (1977). Approaches to Rorschach validation. In M. A. Rickers-Ovsiankina (Ed.), *Rorschach psychology* (pp. 575–608). Huntington, NY: Krieger.
- Weiner, I. B. (1994). The Rorschach Inkblot Method (RIM) is not a test: Implications for theory and practice. *Journal of Personality Assessment, 62*, 498–504.
- Weiner, I. B. (1999). What the Rorschach can do for you: Incremental validity in clinical applications. *Assessment, 6*, 327–339.
- Weiner, I. B. (2001). Advancing the science of psychological assessment: The Rorschach Inkblot Method as exemplar. *Psychological Assessment, 13*, 423–432.
- Weiner, I. B. (2003). *Principles of Rorschach interpretation* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Westen, D. (1998). The scientific legacy of Sigmund Freud: Toward a psychodynamically informed psychological science. *Psychological Bulletin, 124*, 333–371.
- Widiger, T. A., & Schilling, K. M. (1980). Toward a construct validation of the Rorschach. *Journal of Personality Assessment, 44*, 450–459.
- Wiener, D. N. (1948). Subtle and obvious keys for the MMPI. *Journal of Consulting Psychology, 12*, 164–170.
- Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review, 107*, 101–126.
- Wood, J. M., Nezworski, M. T., & Lilienfeld, S. O. (2003). *What's wrong with the Rorschach? Science confronts the controversial inkblot test*. San Francisco: Jossey-Bass.
- Wood, J. M., Nezworski, M. T., & Stejskal, W. J. (1996). The Comprehensive System for the Rorschach: A critical examination. *Psychological Science, 7*, 3–10.
- Wrobel, T. A., & Lachar, D. (1982). Validity of the Wiener subtle and obvious scales for the MMPI: Another example of the importance of inventory-item content. *Journal of Consulting and Clinical Psychology, 50*, 469–470.
- Zajonc, R. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist, 35*, 151–175.