Introduction to Multimethod Clinical Assessment

Robert F. Bornstein and Christopher J. Hopwood

Suppose you developed numbness in your left leg, along with back pain and balance problems—a worrisome situation to say the least. You go to your physician and describe your symptoms; after listening carefully and asking some questions, your physician renders a diagnosis, writes a prescription, and sends you on your way. You're surprised that the doctor didn't order additional tests, but you assume she must know what she is doing, so you let it go.

A week has gone by and the symptoms have not remitted, so you decide to consult a different physician. As you start to explain your problem, the physician stops you. He doesn't want to hear your take on things, he says—tests will provide all the information he needs. He writes out several orders and sends you off to the lab.

These scenarios illustrate a fundamental difference between psychology and medicine. In medicine, multimethod assessment is so firmly entrenched that we hardly notice it: Virtually every clinical decision reflects an integration of the patient's self-report with data obtained in other modalities. A very different situation characterizes contemporary psychology. Here the clinician often relies primarily—sometimes completely—on tests from a single modality (e.g., a series of questionnaires), and neither patient nor referent is likely to question this approach. To be sure, many clinicians draw upon evidence from multiple modalities in formulating an assessment, but

in psychology, unlike medicine, this is neither expected nor required. It depends on the background, attitudes, and preferences of the clinician.

We believe it is time for multimethod clinical assessment to become the norm—an expectation rather than an option. Just as physicians cannot gain complete understanding of a patient's problem unless they integrate evidence from multiple modalities (e.g., self-report, behavioral, physiological), psychologists cannot gain complete understanding of a patient's difficulties without evidence from multiple modalities (e.g., self-report, behavioral, performance-based). Moreover, just as the particular combination of tests that is most useful to an assessing physician will depend on the nature of the patient and that patient's symptoms, the particular combination of tests that will be useful to the assessing psychologist will vary from problem to problem and from patient to patient.

What Is Clinical Assessment?

Psychologists, psychiatrists, and others often use the terms *diagnosis* and *assessment* interchangeably, but in fact these terms mean different things. Diagnosis involves identifying and documenting a patient's symptoms to classify that patient into one or more categories whose labels represent shorthand descriptors of complex psychological syndromes (e.g., bulimia nervosa, schizoid personality disorder). Assessment, in contrast, involves administering a series of psychological tests to disentangle the complex array of dispositional and situational factors that combine to determine a patient's subjective experiences, core beliefs, emotional patterns, motives, traits, defenses, and coping strategies.

As a number of writers have noted, diagnosis is key to understanding a patient's *pathology*; assessment is key to understanding the *person* with this pathology (Finn, 2005; Hopwood, 2010; Weiner, 2000). Although assessment data by themselves are not adequate to render a diagnosis, these data can be useful in refining diagnoses, in supporting a tentative diagnosis, and in making a differential diagnosis (such as when marked discontinuities in a patient's performance on more structured versus less structured tests suggest borderline pathology; see Carr & Goldstein, 1981). Beyond refining diagnostic decisions, assessment data play an important role in risk management (VandeCreek & Knapp, 2000) and treatment planning (Clarkin, 2012; Livesley, 2005), as well as in forensic settings (e.g., custody and competency hearings; Hilsenroth & Stricker, 2004), fitness for duty evaluations (Anfang & Wall, 2006), psychiatric disability evaluations (Gold et al., 2008), and myriad other domains.

As is true for diagnosis and assessment, clinicians and researchers often use the terms *psychological testing* and *psychological assessment* interchangeably, but in fact these terms also mean different things. Handler and Meyer (1998) provided an excellent summary of the conceptual

and practical differences between psychological testing and psychological assessment. They wrote:

> Testing is a relatively straightforward process wherein a particular test is administered to obtain a particular score or two. Subsequently, a descriptive meaning can be applied to the score based on normative, nomothetic findings. . . . Psychological assessment, however, is a quite different enterprise. The focus here is not on obtaining a single score, or even a series of test scores. Rather, the focus is on taking a variety of test-derived pieces of information, obtained from multiple methods of assessment, and placing these data in the context of historical information, referral information, and behavioral observations in order to generate a cohesive and comprehensive understanding of the person being evaluated. (pp. 4–5)

Handler and Meyer's (1998) insightful analysis has been echoed and elaborated by numerous clinicians and clinical researchers (e.g., Groth-Marnat, 1999; Widiger & Samuel, 2005). Cates (1999, p. 637) put it well when he noted that in the realm of psychological assessment, "art rests on science." Psychological testing requires precision, objectivity, and the kind of scientific detachment that facilitates accurate data gathering. Psychological assessment involves integration, synthesis, and clarification of ambiguous—even conflicting—evidence obtained during the testing process. As Bornstein (2010) noted in describing the complexity of clinical assessment:

> The competent tester must be (at least for that moment) a staunch behaviorist, understanding the contingencies that define the testing situation and using this knowledge to maximize the validity and generalizability of test data. Once these data are gathered the behavioral tester must transform into a psychodynamically informed assessor, able to combine dynamic concepts with research findings from other areas of psychology to interpret test results in the context of referral information, life history information, and behavioral observations made during testing. (p. 147)

On Test Score Convergences and Discontinuities

Several decades ago psychological assessment almost invariably included a comprehensive test battery consisting of measures designed to tap different domains of adaptation (e.g., trait scales and intellectual tests), and different levels of functioning and experience (e.g., questionnaires and performance-based measures; see Allison, Blatt, & Zimet, 1968; Rapaport, Gill, & Schafer, 1945, 1968). Owing in part to the demands of managed care (Sperling, Sack, & Field, 2000), assessment now consists primarily of the administration, scoring, and interpretation of questionnaires. This trend extends beyond clinical assessment to research settings as well:

When Bornstein (2003) conducted a systematic survey of personality disorder studies published in five major journals between 1991 and 2000, he found that over 80% of published investigations relied exclusively on self-report data, both in quantifying personality pathology and in measuring its correlates and consequences (only 4% of published studies assessed actual behavior).

Even when test batteries include measures from multiple modalities, they are not always integrated in a way that is maximally heuristic and clinically useful. Following the tradition established by Campbell and Fiske (1959), during the past 50 years psychological assessment research has focused primarily on documenting the convergence of scores on different measures of the same construct, even when the measures use very different methods to quantify these constructs (see Messick, 1989, 1995; Slaney & Maraun, 2008). Most clinicians intuitively value converging results from different tests, in part because converging results are reassuring and increase one's confidence in test-derived clinical predictions. In the early 1990s, psychologists began to write more extensively on the systematic interpretation of test score divergences as well as convergences (e.g., Archer & Krishnamurthy, 1993a, 1993b; Meyer, 1996b, 1997). As Meyer et al. (2001) noted, when different personality assessment tools use different formats and engage different psychological processes in the testee, divergences in scores on these tests can be particularly informative.

Consider, for example, a series of studies wherein Bornstein and his colleagues found that discontinuities between self-report and performance-based dependency test scores provided information regarding personality dynamics that neither test alone could provide (Bornstein, 1998; Bornstein, Bowers, & Bonner, 1996a, 1996b; Bornstein, Rossner, Hill, & Stepanian, 1994). These studies were all based on an often observed pattern: Although many patients obtain consistently high (or consistently low) scores on self-report and performance-based dependency tests (and are therefore classified as *high dependent* or *low dependent*), some patients score high on one type of test but low on the other (see Bornstein, 2002, 2012). Those who obtain high performance-based but low self-report dependency scores have a personality style characterized by *unacknowledged dependency*; those who obtain the reverse pattern—low performance-based but high self-report dependency scores—have a *dependent-self presentation*, exaggerating dependent feelings and urges as a means of obtaining rewards in social and work relationships. Moreover, college students who score high on both self-report and performance-based dependency tests have high levels of dependent personality disorder symptoms, whereas students who obtain high performance-based but low self-report dependency scores tend to have histrionic rather than dependent features (Bornstein, 1998). Similar conclusions have emerged in studies contrasting self-report and performance-based measures of other personality traits, including need for

achievement (McClelland, Koestner, & Weinberger, 1989), power (Koestner, Weinberger, & McClelland, 1991), and intimacy (Craig, Koestner, & Zuroff, 1994).

Process and Outcome in Clinical Assessment

Findings like these point to the importance of considering process as well as outcome in clinical assessment: Only by understanding the psychological processes engaged by different types of tests can test score convergences and divergences be interpreted meaningfully. Although a comprehensive analysis of psychological processes that underlie the broad spectrum of tests in use today has yet to be written, Meyer and Kurtz (2006) and others (e.g., Schultheiss, 2007) contrasted the processes engaged by two of the more widely used types of measures: self-report and performance based. As Bornstein (2009) noted, when people genuinely engage a typical self-report test item (e.g., "I would rather be a follower than a leader," "I often feel depleted"), three processes occur in sequence. First, testees engage in *introspection*, turning their attention inward to determine if the statement captures some aspect of their feelings, thoughts, motives, or behaviors. Second, a *retrospective memory search* occurs, as testees attempt to retrieve instances wherein they experienced or exhibited the response(s) described in the test item. Finally, testees may engage in *deliberate self-presentation*, deciding whether, given the context and setting in which they are being evaluated, it is better to answer honestly or to modify their response to depict themselves in a particular way. Typically, these efforts are aimed at "faking good" (i.e., attempting to portray oneself as healthier than is actually the case) or "faking bad" (attempting to portray oneself as unhealthy and exaggerate pathology), depending on the person's self-presentation goals.

Contrast this set of psychological processes with those that occur as people respond to stimuli from a performance-based measure such as the Rorschach Inkblot Method (RIM). Unlike a self-report test, here the fundamental challenge is to create meaning in a stimulus that can be interpreted in multiple ways. To do this, patients must direct their attention outward (rather than inward) and focus on the stimulus (not the self); they then attribute meaning to the stimulus based on properties of the inkblot and the associations primed by these stimulus properties. Once a series of potential percepts (or *stimulus attributions*) is formed, patients typically sort through these possible responses, selecting some and rejecting others before providing their description (see Exner & Erdberg, 2005; Meyer, Viglione, Mihura, Erard, & Erdberg, 2011).

With this as context, Bornstein (2009, 2011) provided a preliminary process-based classification of widely used psychological tests. In an

updated version of this classification, these tests may be divided into five broad categories, as follows.

Self-Attribution Tests

Self-attribution (or *self-report*) test scores reflect the degree to which the person attributes various traits, feelings, thoughts, motives, behaviors, attitudes, or experiences to him- or herself. Because they are efficient and cost effective, self-attribution tests are far and away the most widely used type of test in both research and clinical settings. The Beck Depression Inventory, the Personality Assessment Inventory, and the NEO Personality Inventory (NEO-PI) would all be included in this category, as would questionnaire measures of attitudes, interests, and values.

Stimulus Attribution Tests

Traditionally called *projective tests*, and more recently *performance-based tests*, in stimulus attribution tests the respondent attributes meaning to an ambiguous stimulus, with attributions determined in part by stimulus characteristics and in part by the person's cognitive style, emotions, motives, and need states. The RIM is the most widely used and well-known stimulus attribution test; others include the Thematic Apperception Test and the Holtzman Inkblot Test.

Constructive Tests

In constructive tests, generation of test responses requires the person to create or construct a novel image or written description within parameters defined by the tester. The Draw a Person Test (and other projective drawings) would be classified in this category, as would various open-ended self-descriptions (e.g., Blatt's Qualitative and Structural Dimensions of Object Representations).

Behavioral Tests

In some behavioral tests, scores are derived from indices of a person's behavior exhibited and measured *in vivo*, as in spot sampling (a technique wherein researchers sample behavior at randomly selected times, in multiple contexts). Behavior may also be examined in a controlled setting (e.g., using joystick feedback tasks wherein moment-by-moment behaviors are rated as they occur). Other behavioral tests assess the person's unrehearsed performance on one or more structured tasks designed to tap attentional resources, working memory, and other cognitive skills (e.g., the Bender Visual–Motor Gestalt Test, the Attentional Capacity Test).

Informant-Report Tests

Scores on tests in this category are based on informants' ratings or judgments of a person's characteristic patterns of responding (e.g., the therapist version of the Shedler–Westen Assessment Procedure, the Informant Report version of the NEO-PI). In contrast to observational measures, which are based on direct observation of behavior, informant-report tests are based on informants' retrospective, memory-derived conclusions regarding characteristics of the target person.[1]

Dispositional and Contextual Influences on Psychological Test Responses

Scrutiny of these five categories and the processes engaged by tests within each category suggests two things. First, scores on each type of test not only reflect aspects of the construct that the test is designed to assess (e.g., narcissism, self-esteem, introjects, cognitive skills), but are also influenced by an array of dispositional and contextual variables, not all of which are conceptually linked with the construct in question. Second, in many instances these extraneous variables—typically considered confounds in clinical assessment—will differentially influence scores derived from tests in different categories, even when these tests purport to quantify the same construct.

Among the key dispositional influences on psychological test responses are self-perception biases (i.e., distortions in the respondent's view of him- or herself; Oltmanns & Turkheimer, 2009), and memory distortions (e.g., selective recall of trait-relevant behaviors). The individual's cognitive (information-processing) style also plays a role: Some people tend to focus primarily on details when thinking about themselves, other people, or test stimuli (e.g., inkblots); others emphasize overall patterns and global impressions at the expense of detail. Self-attributions and informant reports are both influenced by various heuristics inherent in self-perception and perceptions of others (e.g., confirmatory bias, actor–observer effects, the fundamental attribution error). Finally, studies confirm that self-presentation needs (e.g., the desire to present oneself in a positive or negative light) often influence psychological test responses (Horvath & Morf, 2010), as do the

[1]Certain psychological tests have characteristics of more than one category and may be best conceptualized as "hybrid tests" that engage multiple processes. For example, structured clinical interviews involve patient self-report (and therefore engage the processes characteristic of self-attribution tests), as well as clinician observation of patient behavior (and therefore engage the processes of behavioral tests). Broad-range neuropsychological tests like the Halstead–Reitan may engage processes from multiple categories (see Reitan & Wolfson, 2000).

person's previous testing experiences and the expectations that these earlier experiences produce (Garb, 1998).

Beyond dispositional influences, several state and contextual variables have been shown to play a role in moderating test scores. For example, variations in the respondent's mood or anxiety level influence test responses in at least two ways. First, variations in mood influence retrieval of episodic memories, as people tend to retrieve mood-congruent memories more readily than mood-incongruent ones (Rholes, Riskind, & Lane, 1987). Second, anxiety captures attentional capacity (as does negative mood), impairing respondents' performance on various behavioral and cognitive measures (Arnell, Killman, & Fijavz, 2007). In addition, sometimes particular concepts or motives are inadvertently primed during a testing situation, influencing test performance: Masling and others have shown how the gender of the examiner influences RIM responses (see Masling, 1966, 2002, for reviews); even seemingly minor stimuli such as an examiner's clothing, his or her age, and the layout of the testing room may affect test performance (Bargh & Morsella, 2008; Weiner, 2004).

Quantifying Dispositional and Contextual Influences

There are at least three ways of assessing the impact of various dispositional and contextual variables on psychological test scores. First, researchers can examine naturally occurring (*in vivo*) influences (e.g., variations in mood or in anxiety level). This was the approach used by Hirschfeld, Klerman, Clayton, and Keller (1983) to examine the impact of changes in severity of depressive symptoms on traits theoretically linked with depression (e.g., dependency, self-esteem). Second, researchers can examine changes in test scores over time due to the effects of maturation (in children) or aging (in older adults). This was the approach used by Jansen and Van der Maas (2002) to detect Piagetian developmental shifts in children's reasoning, and that used by Baltes (1996) to assess age-related changes in the expression of underlying dependency needs in older adults (see also Roberts & DelVecchio, 2000, for additional findings in this area).

A third approach—the least widely used but potentially the most informative—is to introduce experimental manipulations that deliberately alter the processes engaged by different psychological tests. This approach allows the researcher to (1) confirm that altering these processes does in fact change test scores as expected; and (2) illuminate the processes involved in two tests that measure parallel constructs using contrasting methods. This was the approach used by Bornstein et al. (1994, 1996a) to examine the differential impact of instructional set and mood state on self-report and performance-based dependency scores. As hypothesized, deliberately inducing a negative mood state increased performance-based (but not self-report) dependency scores, whereas an instructional set that

framed interpersonal dependency in negative terms increased self-report (but not performance-based) dependency. Along somewhat similar lines, Morf and her colleagues have examined the impact of threats to self-esteem on self- and other-evaluations in narcissistic and control participants (see Morf & Rhodewalt, 2001). Arntz and his colleagues assessed the impact of manipulating stress level and mood on schema-related responding in patients with and without borderline pathology (e.g., Arntz et al., 2009).

Maximizing the Value of Multimethod Assessment: A Framework for Test Score Integration

In many—perhaps most—testing situations, multimethod assessment will yield richer, more clinically useful data than assessment that relies exclusively on tests from a single modality. From a psychometric standpoint, multimethod assessment helps minimize the negative impact of reliability and validity limitations inherent in different types of measures, because these limitations tend to vary across test modality. Although they can never be entirely eliminated, to some degree these limitations can be balanced out by deliberately selecting tests with contrasting strengths and weaknesses. From a clinical standpoint, when test data from different modalities are integrated, and test score convergences and divergences are explored, multimethod assessment allows aspects of a patient's dynamics that might otherwise go unrecognized to be scrutinized directly (e.g., conflicts, defenses, unconscious motives, and emotional responses, areas wherein the patient has limited insight or is overtly self-deceptive).

Thus, in our view, the central issue regarding multimethod clinical assessment is not *why* but *how*. In the following sections we outline a six-step framework for multimethod assessment and test score integration.

1. *Understand the strengths and limitations of different methods*. In part, these methods reflect the psychometric properties of each measure (see Messick, 1989, 1995) and the degree to which scores derived from that measure fulfill established criteria for validity (convergent, discriminant, concurrent, predictive), and reliability (retest, internal, interrater). The strengths and limitations of different methods are also a product of the psychological processes engaged by measures within that particular test category (Bornstein, 2011), since different processes (e.g., self-attributions, online responding, judgments of others' behavior) are differentially influenced by various extraneous variables (e.g., self-presentation goals, testing milieu).

2. *Know when to collect data using multiple methods*. Although critics of the RIM have often cited its modest correlations with self-report test scores as evidence of poor RIM validity (e.g., Wood, Nezworski,

Lilienfield, & Garb, 2003), given the different processes engaged by self-report and performance-based tests, such modest correlations—far from being problematic—actually represent evidence supporting the discriminant validity of both measures (Bornstein, 2002; McGrath, 2008). Similarly, while modest correlations between scores on the NEO-PI and indices of observable behavior have occasionally been cited as evidence of limitations in the measure (Block, 2010), given respondents' self-presentation needs and inherent limitations in our ability to describe ourselves accurately, correlations (effect sizes) between self-report test scores and behavior in the medium range are precisely what one would expect for trait-focused scales.

Any time two tests that measure parallel constructs using different methodologies fulfill established criteria for reliability and validity, each test has the potential to add incremental validity—unique predictive value—to a test battery. Thus, use of two tests that measure a particular construct via contrasting methods is potentially useful any time a complete and nuanced understanding of this construct is needed (see Meyer et al., 2001). Given cost and efficiency concerns, multimethod assessment of a construct is most easily justified when a complex clinical or empirical question merits particularly close scrutiny (e.g., when assessing impulse control, suicidality, parental fitness, or competence to stand trial).

3. *Decide which methods to use.* The choice of assessment method will be based in part on the referral question, the patient's history, and results from previous evaluations if these are available. The domains of behavior and mental functioning most salient to the assessment (e.g., stress tolerance, potential to benefit from psychotherapy) are also relevant here: Because different measures are best suited for predicting different forms of behavior, it is important to tailor the battery to match tests with outcome. (See, e.g., findings demonstrating that performance-based measures of interpersonal dependency and need for achievement predict spontaneous behavior in those domains, whereas self-report measures of these constructs tend to predict goal-directed rather than spontaneous behavior; Bornstein, 2002; McClelland et al., 1989.) Meyer (1996a) suggested that—even within a category—psychological tests can be distinguished with respect to the degree of *conscious penetration* associated with the processes engaged by that test (i.e., the degree to which test responses reflect deliberate, mindful responding versus reflexive, automatic processing). Erdelyi (2004) has shown that conscious awareness of internal states may also vary over time, waxing and waning in response to external events and environmental contingencies.

4. *Select appropriate measures.* Beyond the referral question itself, initial test selection decisions should be based on validity evidence, an understanding of underlying processes engaged by different tests, cost effectiveness, and clinical utility. Given that clinical assessment is a dynamic process

(Finn, 2005; Hopwood, 2010; Hopwood & Huprich, 2011), and that new questions arise as data accumulate, in many instances the assessor must adjust "on the fly," modifying the test battery as preliminary results reveal new issues that merit scrutiny. For example, a patient may produce an unremarkable Minnesota Multiphasic Personality Inventory (MMPI) clinical scale profile, but if scrutiny of that patient's validity scales suggests defensive responding, it may be useful to follow up with a performance-based test that includes well-validated indices of psychopathology and subjective distress. In general, scales with high face validity are more susceptible to dissimulation and self-presentation effects than are scales with low face validity (Shedler, Mayman, & Manis, 1993); as Bornstein et al. (1994) and McGrath (2008) noted, performance based tests in general tend to have low face validity. Self-report tests also vary with respect to face validity, however, and even within a particular scale items may vary considerably with respect to susceptibility to dissimulation and self-presentation effects (Sartori, 2010).

5. *Implement a framework for integrating data from different sources.* How can we integrate data we acquire from multiple sources? First, we must develop an overarching framework for understanding convergences and divergences among scores from tests that measure parallel constructs using different methods (see Finn, 2007). Bornstein's (1998, 2012) four-cell model of interpersonal dependency based on the integration of self-report and performance-based test patterns may be useful in this context; in this model respondents are classified into *low dependent*, *high dependent*, *unacknowledged dependency*, and *dependent self-presentation groups* (see Bornstein, 2012, Fig. 1). A parallel framework was developed by Shedler et al. (1993) when they used responses to the Eysenck Neuroticism Scale and raters' evaluations of open-ended descriptions of early memories to classify respondents into *genuinely healthy*, *genuinely distressed*, *"illusion of mental illness"* (exaggerations in self-reported distress), and *"illusion of mental health"* (defensively healthy) groups (see Shedler et al., 1993, Fig. 1).

Second, we must develop a framework for integrating test data across domains as well as across methods. Carr and Goldstein's (1981) finding that marked discontinuities in patient performance between more structured versus less structured tests is useful in identifying underlying borderline pathology is an example of this strategy (see also Hopwood et al., 2008, for evidence regarding the contrasting dynamics of questionnaires and structured interviews). This is also the general approach used by neuropsychologists to identify areas of cognitive deficit by contrasting patients' performance across tests that capture different skills and capacities. Along slightly different lines, Hopwood et al. (2011) contrasted circumplex-derived indices of interpersonal sensitivities across three types of relationships (romantic, platonic, and non-close) to elucidate the degree to which consistent interpersonal patterns emerged across relationship domains.

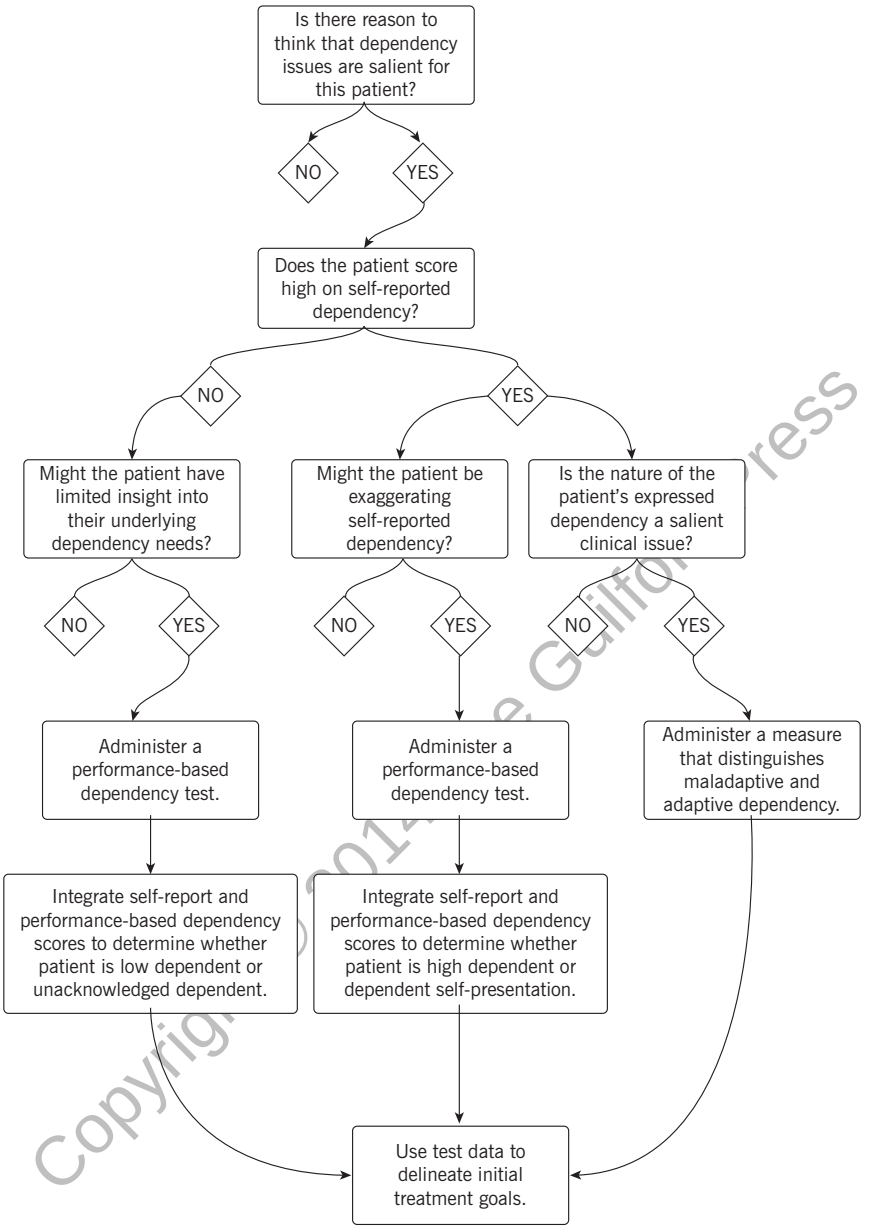

FIGURE 1. Decision tree for using multimethod dependency test data in treatment planning. Initial assessment of the patient's self-reported (self-attributed) dependency may be followed by administration of a performance-based dependency test, if there is reason to believe the patient may be either underreporting or exaggerating his or her dependency. If the nature of the patient's expressed dependency is a salient clinical issue, then initial assessment of self-reported dependency may be followed by administration of a measure that distinguishes maladaptive from adaptive expressions of underlying dependency needs.

Similar logic holds when integrating personality and psychopathology test data with data derived from a patient's life records, and with that provided by knowledgeable informants; in both areas convergences and discontinuities may be informative.

6. *Use assessment data to enhance treatment planning.* Moving from assessment to treatment planning is a stepwise process: The initial clinical (or referral) question will guide the clinician's preliminary choice of tests, after which the initial test results—interpreted in the context of the patient's life history, presenting problem, and other salient information—will determine the next test(s) to be administered, and how test results may be integrated most usefully. Figure 1 summarizes this process, illustrating how initial test results guide subsequent clinical decisions when assessing patient dependency.

Multimethod Clinical Assessment: Looking Forward

A plethora of empirical research indicates that use of multiple assessment methods in clinical and research settings provides important incremental information that cannot be obtained when a single assessment modality is used. Nevertheless, many psychologists continue to utilize unimodal rather than multimethod assessment in their practice and research, in part because empirically validated, clinically useful models for integrating multimethod data have not been presented in a comprehensive, systematic, transtheoretical way. The goal of this volume is to strengthen links between evidence-based multimethod assessment and clinical practice by providing systematic reviews of how to incorporate diverse assessment techniques in the laboratory, clinic, and consulting room.

The volume includes 14 chapters by leading clinical researchers. Within each chapter we asked authors to (1) discuss the assessment approaches that are particularly useful for assessing key constructs relevant to a particular clinical issue, along with a rationale for integrating data within this domain; (2) review empirical evidence supporting the integration of these methods, including evidence regarding their interrelations and the incremental validity provided by each method; and (3) describe a case in which assessment using these assessment methods was clinically useful. Chapters are organized into three broad domains, as follows.

The first five chapters focus on *Personality and Individual Differences*, opening with a discussion of multimethod assessment of personality traits by Galione and Oltmanns. Pincus, Sadler, Woody, Roche, Thomas, and Wright review research on multimethod assessment of interpersonal dynamics, followed by Tomko and Trull's review studies of affective processes. Kosloff, Maxfield, and Solomon discuss existential concerns. This section closes with Cogswell and Emmert's discussion of implicit processes,

one of the most challenging areas in contemporary psychological assessment and one that helped impel many of the current integrative approaches used in this area.

The next four chapters discuss multimethod assessment of *Psychopathology and Resilience*. Here authors address the opportunities and challenges faced in assessing several variables of particular relevance in the clinical setting: anxiety (Moser, Przeworski, Schroder, & Dunbeck), externalizing disorders (Blonigen & Wytiaz), thought quality (Blais & Bello), and resilience (Denckla & Mancini). Chapters in this section illustrate nicely the range of assessment modalities used to assess psychopathology and protective factors, and the contrasting strategies that are useful in integrating assessment data when different constructs are examined.

The final section includes chapters on *Clinical Management*. Mihura and Graceffo review research on multimethod assessment and treatment planning, followed by Pascual-Leone, Singh, Harrington, and Yeryomenko's discussion of how assessment can illuminate the treatment process. Burchett and Bagby review research on detecting and evaluating distortion and dissimulation in patient responding, followed by Stanfill, O'Brien, and Viglione's discussion of the complexities of multimethod risk assessment. The volume closes with Smith and Finn's review of the therapeutic presentation of multimethod assessment results, which brings us full circle, as Smith and Finn demonstrate how assessment data not only inform the clinician, but may also facilitate patient motivation and engagement.

As these brief descriptions illustrate, this volume addresses a diverse array of clinical issues. A broad range of methods are used to address these various issues, including self-reports, performance-based tests, behavioral measures (both laboratory-based and *in vivo*), archival data, and observer reports. Despite this diversity in topic, method, and integration strategy, contributors share a fundamental belief in the value of multimethod assessment in contemporary clinical psychology. As a result, these chapters not only bring together the best work on multimethod clinical assessment available today, but help set the stage for continued refinement of empirically validated integrative assessment methods that will enhance this important area of clinical practice and research during the coming years.

REFERENCES

Allison, J., Blatt, S. J., & Zimet, C. N. (1968). *The interpretation of psychological tests*. New York: Harper & Row.

Anfang, S. A., & Wall, B. W. (2006). Psychiatric fitness for duty evaluations. *Psychiatric Clinics of North America*, *29*, 675–693.

Archer, R. P., & Krishnamurthy, R. (1993a). Combining the Rorschach and 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.

Arnell, K. M., Killman, K. V., & Fijavz, D. (2007). Blinded by emotion: Target misses follow attention capture by arousing distractors in RSVP. *Emotion*, *7*, 465–477.

Arntz, A., Roos, D., & Dreesen, L. (2009). Assumptions in borderline personality disorder: Specificity, stability, and relationship with etiological factors. *Behaviour Research and Therapy*, *37*, 545–557.

Baltes, M. M. (1996). *The many faces of dependency in old age*. Cambridge, UK: Cambridge University Press.

Bargh, J. A., & Morsella, E. (2008). The unconscious mind. *Perspectives on Psychological Science*, *3*, 73–79.

Block, J. (2010). The five-factor framing of personality and beyond: Some ruminations. *Psychological Inquiry*, *21*, 2–25.

Bornstein, R. F. (1998). Implicit and self-attributed dependency needs in dependent and histrionic personality disorders. *Journal of Personality Assessment*, *71*, 1–14.

Bornstein, R. F. (2002). A process dissociation approach to objective–projective test score interrelationships. *Journal of Personality Assessment*, *78*, 47–68.

Bornstein, R. F. (2003). Behaviorally referenced experimentation and symptom validation: A paradigm for 21st century personality disorder research. *Journal of Personality Disorders*, *17*, 1–18.

Bornstein, R. F. (2009). Heisenberg, Kandinsky, and the heteromethod convergence problem: Lessons from within and beyond psychology. *Journal of Personality Assessment*, *91*, 1–8.

Bornstein, R. F. (2010). Psychoanalytic theory as a unifying framework for 21st century personality assessment. *Psychoanalytic Psychology*, *27*, 133–152.

Bornstein, R. F. (2011). Toward a process-focused model of test score validity: Improving psychological assessment in science and practice. *Psychological Assessment*, *23*, 532–544.

Bornstein, R. F. (2012). From dysfunction to adaptation: An interactionist model of dependency. *Annual Review of Clinical Psychology*, *8*, 291–316.

Bornstein, R. F., Bowers, K. S., & Bonner, S. (1996a). Effects of induced mood states on objective and projective dependency scores. *Journal of Personality Assessment*, *67*, 324–340.

Bornstein, R. F., Bowers, K. S., & Bonner, S. (1996b). Relationships of objective and projective dependency scores to sex role orientation in college students. *Journal of Personality Assessment*, *66*, 555–568.

Bornstein, R. F., Rossner, S. C., Hill, E. L., & Stepanian, M. L. (1994). Face validity and fakability of objective and projective measures of dependency. *Journal of Personality Assessment*, *63*, 363–386.

Campbell, D. T., & Fiske, D. (1959). Convergent and discriminant validation by the multitrait–multimethod matrix. *Psychological Bulletin*, *56*, 81–105.

Carr, A. C., & Goldstein, E. G. (1981). Approaches to the diagnosis of borderline conditions by the use of psychological tests. *Journal of Personality Assessment*, *45*, 563–574.

Cates, J. A. (1999). The art of assessment in psychology: Ethics, expertise, and validity. *Journal of Clinical Psychology*, *55*, 631–641.

Clarkin, J. F. (2012). An integrated approach to psychotherapy techniques for patients with personality disorder. *Journal of Personality Disorders*, *26*, 43–62.

Craig, J. A., Koestner, R., & Zuroff, D. C. (1994). Implicit and self-attributed intimacy motivation. *Journal of Social and Personal Relationships*, *11*, 491–507.

Erdelyi, M. H. (2004). Subliminal perception and its cognates: Theory, indeterminacy, and time. *Consciousness and Cognition*, *13*, 73–91.

Exner, J. E., & Erdberg, P. (2005). *The Rorschach: Advanced interpretation* (3rd ed.). New York: Wiley.

Finn, S. E. (2005). How psychological assessment taught me compassion and firmness. *Journal of Personality Assessment*, *84*, 29–32.

Finn, S. E. (2007). *In our client's shoes: Theory and techniques of therapeutic assessment*. New York: Routledge/Taylor & Francis.

Garb, H. N. (1998). *Studying the clinician: Judgment research and psychological assessment*. Washington, DC: American Psychological Association.

Gold, L. H., Anfang, S. A., Drukteinis, A. M., Metzner, J. L., Price, M., Wall, B. W., et al. (2008). AAPL practice guideline for the forensic evaluation of psychiatric disability. *Journal of the American Academy of Psychiatry and the Law*, *36*, S3–S50.

Groth-Marnat, G. (1999). Current status and future directions of psychological assessment. *Journal of Clinical Psychology*, *55*, 781–785.

Handler, L., & Meyer, G. J. (1998). The importance of teaching and learning personality assessment. In L. Handler & M. J. Hilsenroth (Eds.), *Teaching and learning personality assessment* (pp. 3–30). Mahwah, NJ: Erlbaum.

Hilsenroth, M. J., & Stricker, G. (2004). A consideration of challenges to psychological assessment instruments used in forensic settings: Rorschach as exemplar. *Journal of Personality Assessment*, *83*, 141–152.

Hirschfeld, R. M. A., Klerman, G. L., Clayton, P. J., & Keller, M. B. (1983). Personality and depression: Empirical findings. *Archives of General Psychiatry*, *40*, 993–998.

Hopwood, C. J. (2010). An interpersonal perspective on the personality assessment process. *Journal of Personality Assessment*, *92*, 471–479.

Hopwood, C. J., Ansell, E. B., Pincus, A. L., Wright, A. G. C., Lukowitsky, M. R., & Roche, M. J. (2011). The circumplex structure of interpersonal sensitivities. *Journal of Personality*, *79*, 707–739.

Hopwood, C. J., & Huprich, S. K. (2011). Introduction to the Special Issue on personality assessment in the DSM-5. *Journal of Personality Assessment*, *93*, 323–324.

Hopwood, C. J., Morey, L. C., Edelen, M. O., Shea, M. T., Grilo, C. M., Sanislow, C. A., et al. (2008). A comparison of interview and self-report methods for the assessment of borderline personality disorder criteria. *Psychological Assessment*, *20*, 81–85.

Horvath, S., & Morf, C. C. (2010). To be grandiose or not to be worthless: Different routes to self-enhancement for narcissism and self-esteem. *Journal of Research in Personality*, *44*, 585–592.

Jansen, B. R. J., & Van der Maas, H. L. J. (2002). The development of children's rule use on the balance scale task. *Journal of Experimental Child Psychology*, *81*, 383–416.

Koestner, R., Weinberger, J., & McClelland, D. C. (1991). Task-intrinsic and social-extrinsic sources of arousal for motives assessed in fantasy and self-report. *Journal of Personality*, *59*, 57–82.

Livesley, W. J. (2005). Principles and strategies for treating personality disorder. *Canadian Journal of Psychiatry*, *50*, 442–450.

Masling, J. M. (1966). Role-related behavior of the subject and psychologist and its effect upon psychological data. In D. Levine (Ed.), *Nebraska symposium on motivation* (pp. 67–104). Lincoln: University of Nebraska Press.

Masling, J. M. (2002). Speak, memory, or goodbye Columbus. *Journal of Personality Assessment*, *78*, 4–30.

McClelland, D. C., Koestner, R., & Weinberger, J. (1989). How do self-attributed and implicit motives differ? *Psychological Review*, *96*, 690–702.

McGrath, R. E. (2008). The Rorschach in the context of performance based personality assessment. *Journal of Personality Assessment*, *90*, 465–475.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–103). New York: Macmillan.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, *50*, 741–749.

Meyer, G. J. (1996a). 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. (1996b). The Rorschach and MMPI: Toward a more scientific understanding of cross-method assessment. *Journal of Personality Assessment*, *67*, 558–578.

Meyer, G. J. (1997). On the integration of personality assessment methods: The Rorschach and MMPI. *Journal of Personality Assessment*, *68*, 297–330.

Meyer, G. J. (2000). Incremental validity of the Rorschach Prognostic Rating Scale over MMPI Ego Strength Scale and IQ. *Journal of Personality Assessment*, *74*, 356–370.

Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland, K. L., Dies, R. R., et al. (2001). Psychological testing and psychological assessment: A review of evidence and issues. *American Psychologist*, *56*, 128–165.

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., Viglione, D. J., Mihura, J. L., Erard, R. E., & Erdberg, P. (2011). *Rorschach Performance Assessment System: Administration, coding, interpretation, and technical manual*. Toledo, OH: Rorschach Performance Assessment System.

Morf, C. C., & Rhodewalt, F. (2001). Unraveling the paradoxes of narcissism: A dynamic self-regulatory processing model. *Psychological Inquiry*, *12*, 177–196.

Oltmanns, T. F., & Turkheimer, E. (2009). Person perception and personality pathology. *Current Directions in Psychological Science*, *18*, 32–36.

Rapaport, D., Gill, M. M., & Schafer, R. (1945). *Diagnostic psychological testing*. Chicago: Yearbook.

Rapaport, D., Gill, M. M., & Schafer, R. (1968). *Diagnostic psychological testing* (rev. ed.). New York: International Universities Press.

Reitan, R. M., & Wolfson, D. (2000). The neuropsychological similarities of mild and more severe head injury. *Archives of Clinical Neuropsychology*, *15*, 433–442.

Rholes, W. S., Riskind, J. H., & Lane, J. W. (1987). Emotional states and memory biases: Effects of cognitive priming and mood. *Journal of Personality and Social Psychology*, *52*, 91–99.

Roberts, B. W., & DelVecchio, W. F. (2000). The rank-order consistency of personality from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin*, *126*, 3–25.

Sartori, R. (2010). Face validity in personality tests: Psychometric instruments and projective techniques in comparison. *Quality and Quantity*, *44*, 749–759.

Schultheiss, O. C. (2007). A memory systems approach to the classification of personality tests: A response to Meyer and Kurtz (2006). *Journal of Personality Assessment*, *89*, 197–201.

Shedler, J., Mayman, M., & Manis, M. (1993). The illusion of mental health. *American Psychologist*, *48*, 1117–1131.

Slaney, K. L., & Maraun, M. D. (2008). A proposed framework for conducting data-based test analysis. *Psychological Methods*, *13*, 376–390.

Sperling, M. B., Sack, A., & Field, C. L. (2000). *Psychodynamic practice in a managed care environment*. New York: Guilford Press.

VandeCreek, L., & Knapp, S. (2000). Risk management and life threatening patient behaviors. *Journal of Clinical Psychology*, *56*, 1335–1351.

Weiner, I. B. (2000). Using the Rorschach properly in practice and research. *Journal of Clinical Psychology*, *56*, 435–438.

Weiner, I. B. (2004). Rorschach Inkblot Method. In M. E. Maruish (Ed.), *The use of psychological testing for treatment planning and outcomes assessment* (3rd ed., pp. 553–588). Mahwah, NJ: Erlbaum.

Widiger, T. A., & Samuel, D. B. (2005). Evidence-based assessment of personality disorders. *Psychological Assessment*, *17*, 278–287.

Wood, J. M., Nezworski, M. T., Lilienfeld, S. O., & Garb, H. N. (2003). *What's wrong with the Rorschach?* San Francisco: Jossey-Bass.

Guilford Publications
72 Spring Street
New York, NY 10012
212-431-9800
800-365-7006
www.guilford.com