

## Empirical Support for Psychological Assessment in Clinical Health Care Settings

Tom W. Kubiszyn  
American Psychological Association

Gregory J. Meyer  
University of Alaska Anchorage

Stephen E. Finn  
Center for Therapeutic Assessment

Lorraine D. Eyde  
U.S. Office of Personnel Management

Gary G. Kay  
Georgetown University Medical Center

Kevin L. Moreland  
Fort Walton Beach, Florida

Robert R. Dies  
New Port Richey, Florida

Elena J. Eisman  
Massachusetts Psychological Association

Psychologists in health care settings today find it increasingly difficult to obtain authorization and appropriate reimbursement for psychological assessments from 3rd party payers. Authorization and reimbursement denials often are based on allegations that empirical support for the utility and validity of psychological tests is nonexistent or limited. This article reviews a sample of the considerable empirical support that exists for the utility and validity of a variety of psychological tests for a wide range of clinical health care applications. Informed by data such as these, psychologists should become more effective in their abilities to seek authorization and reimbursement for assessment and overturning denials.

Obtaining authorization and reimbursement for psychological assessments from third party payers has become increasingly difficult for psychologists (Eisman et al., 2000; Piotrowski, Belter, & Keller, 1998). Difficulty obtaining preauthorization, limitations in instrument selection and time, and reduced or nonexistent remuneration for psychological assessments have become the norm. As a result, psychologists now use psychological assessment less

frequently than in the past (Eisman et al., 2000; Piotrowski et al., 1998).

Eisman et al. (2000) described a variety of reasons for these developments. One is the erroneous belief among third party payers and others that psychological assessment lacks empirical support for its validity and utility in clinical health care settings and provides nothing more than can be gained by a clinical

---

TOM W. KUBISZYN received his PhD in school psychology from the University of Texas at Austin. He is an adjunct associate professor of educational psychology at the University of Texas at Austin and was an assistant executive director for professional practice with the American Psychological Association (APA) when this article was written.

GREGORY J. MEYER received his PhD in clinical psychology from Loyola University Chicago. He is an associate professor of psychology at the University of Alaska Anchorage.

STEPHEN E. FINN received his PhD in clinical psychology from the University of Minnesota. He is the director of the Center for Therapeutic Assessment in Austin, TX.

LORRAINE D. EYDE received her PhD in industrial/organizational psychology from Ohio State University. She is a personnel research psychologist with the Office of Personnel Management in Washington, DC.

GARY G. KAY received his PhD in clinical psychology from Memphis State University. He is a clinical professor of neuropsychology at Georgetown University Medical Center in Washington, DC.

KEVIN L. MORELAND received his PhD in clinical psychology from the University of North Carolina at Chapel Hill. He was a consultant in Ft. Walton Beach, FL.

ROBERT R. DIES received his PhD in clinical psychology from the University of Connecticut. He maintains an independent practice in New Port Richey, FL.

ELENA J. EISMAN received her PhD in counseling psychology from Boston University. She is the executive director of the Massachusetts Psychological Association.

A VERSION OF THIS ARTICLE was initially prepared by the Psychological Assessment Work Group (PAWG) established by the Board of Professional Affairs of the APA. The work group was chaired by Stephen E. Finn. We are grateful to the many APA staff members who provided their support and assistance to the work group and to the Society for Personality Assessment and Rorschach Workshops for supporting Gregory J. Meyer to review much of the literature discussed here; to Stephen N. Haynes, who helped identify some of the relevant literature; and to Muriel D. Lezak for her helpful advice. We also were saddened to learn of the death of Kevin L. Moreland in August 1999. His contributions to the work group and the field of psychological assessment were significant, and he will be missed. THE OPINIONS EXPRESSED IN THIS ARTICLE ARE THOSE OF THE AUTHORS AND DO NOT NECESSARILY REFLECT THE OFFICIAL POLICY OF THE U.S. OFFICE OF PERSONNEL MANAGEMENT.

CORRESPONDENCE CONCERNING THIS ARTICLE should be addressed to Tom W. Kubiszyn, Department of Educational Psychology, Sanchez Building 504, University of Texas, Austin, Texas 78712. Electronic mail may be sent to tom.kubiszyn@mail.utexas.edu.

interview alone. For example, Ambrose (1997), the clinical director of a large managed care organization (MCO), concluded that although psychologists claim that psychological assessment improves diagnosis and treatment outcomes and shortens treatment, "there is no conclusive, unequivocal research that demonstrates assessment does any of the above" (p. 66). Anecdotal information collected from practicing psychologists revealed that Ambrose's position is representative of other MCO clinical staff and reviewers (Eisman et al., 2000). Thus it is no surprise that this is one argument offered by third-party payers to limit authorization, instrument selection, time, and reimbursement for psychological assessment.

If empirical support for psychological assessment were lacking we would have little recourse other than to accept this state of affairs. To the contrary, considerable empirical support exists for many important clinical health care applications of psychological assessment instruments. For such applications psychological assessment can enhance diagnosis and treatment. Health care costs savings would be expected to follow from enhanced diagnosis and treatment, an outcome that third-party payers would be expected to be seriously interested in. Yet, what psychological assessment has to offer has not been heard or understood by third-party payers and other decision makers, and even by some psychologists. However, if the research support exists, why has it not been brought to bear effectively to counter the recent decline of psychological assessment?

The answer to this question may lie in the sheer volume, breadth, and technical complexity of the body of research related to psychological assessment that has evolved over several decades. This body of research is now so large that no practitioner could be expected to take or have the time to identify, collect, read, understand, and summarize this highly technical evidence while also conducting a clinical practice.

The American Psychological Association's Board of Professional Affairs (BPA) established a Psychological Assessment Work Group (PAWG) in 1996 and charged it "with two tasks: 1) to assess the scope of the threat to psychological and neuropsychological assessment services in the current healthcare delivery system, and 2) to assemble key pieces of research that demonstrate the efficacy of psychological assessment services in clinical practice." The PAWG's findings and recommendations were released in two reports to the BPA in 1998 (Eisman et al., 1998; Meyer et al., 1998).<sup>1</sup>

The article by Eisman et al. (2000) is based on the second report. It documents the decline and the reasons for it, and it recommends actions to arrest the decline. This article and a companion article (Meyer et al., 1999) are based on the first PAWG report. The Meyer et al. article (a) compared the validity of psychological assessment with common medical and dental tests and treatments, (b) documented empirically the limitations of clinical interviews alone for diagnostic purposes and explained why improved validity may be expected from multimethod psychological assessments, (c) clarified the important but often misunderstood distinction between psychological assessment and psychological testing, and (d) outlined needed future research directions.

This article complements the Meyer et al. (1999) article by identifying some of the relevant research that demonstrates the validity and utility of psychological assessment for several clinical health care applications. These applications include the (a) descrip-

tion of clinical symptomatology and differential diagnosis; (b) description and prediction of functional behavior; (c) prediction of health outcomes; (d) prediction of health care utilization; (e) prediction of psychotherapy, forensic, and mental health outcomes; (f) identification of patient characteristics that affect treatment; and (g) use of psychological assessment as treatment in itself. These applications were those judged by the PAWG to be among the best supported and most important uses for psychological assessment today. The list is not intended to be exhaustive. Exclusion of other applications does not imply that psychological assessment is not valid for uses not reviewed, or that other uses are unnecessary or unimportant.

Next, we consider the evidence supporting the utility of psychological assessment for each of these applications. When appropriate, we review for each application the data that support the validity and utility of different categories of psychological assessment methods (e.g., neuropsychological tests, self-report tests, and performance-based personality tests).<sup>2</sup> Whenever available, meta-analyses are considered first, followed by large-scale and then small-scale studies.

We hope that this review will serve as a user-friendly resource for assessment-minded psychologists. We hope that these data, coupled with the data reported by Meyer et al. (1999), will enable practitioners to be better prepared and informed in seeking authorization or reimbursement for psychological assessment from third-party payers, and in appealing denials.

## Description of Clinical Symptomatology and Differential Diagnosis

Neuropsychological tests, self-report questionnaires, and performance-based personality tests all have demonstrated the ability to identify, describe, and quantify important patient characteristics. Thus, these psychological tests are useful both for describing clinical symptoms and syndromes that may be targets for treatment and for facilitating accurate diagnosis toward cost-effective treatment in both mental health and medical settings.

### Neuropsychological Tests

In a meta-analysis of 77 studies, Christensen, Hadzi-Pavlovic, and Jacomb (1991) found that neuropsychological tests were ef-

<sup>1</sup> Both PAWG reports can be obtained from Christopher J. McLaughlin, Assistant Director for Professional Development, Practice Directorate, American Psychological Association, 750 First Street, N.E., Washington, DC, 20002-4242; electronic mail may be sent to cmclaughlin@apa.org.

<sup>2</sup> We use the phrase "performance-based personality test" to refer to tasks such as the Rorschach and Thematic Apperception Test (TAT). Historically, these tests have been termed *projective tests*. We use the alternative terminology to emphasize two points. First, much like cognitive and neuropsychological tests, these tasks require patients to perform a defined activity with an examiner (i.e., generate a story or identify images). Second, evidence clearly indicates that these tests do not depend on or require a process of "projection," as it has historically been defined (Exner, 1989). More often than not, the information obtained from these tasks reflects the patient's perceptions, classifications, and cognitive-emotional templates or internal representations, rather than projections per se.

fective (i.e., produced very large effect sizes;  $r = .68$ )<sup>3</sup> in differentiating the normal elderly person from patients with mild, moderate, or severe dementia. Meta-analytic and narrative reviews have also shown that neuropsychologists make reliable and accurate judgments when they use a battery of test data to make inferences about cognitive impairment due to brain damage (Garb & Schramke, 1996; Russell, 1995). Similar accuracy cannot be obtained simply through interviews or informal observation (e.g., Roca, Klein, & Vogelsang, 1982; Schwartz & Wiedel, 1981). In a meta-analysis of 39 studies, Parker, Hanson, and Hunsley (1988) demonstrated that the Wechsler Adult Intelligence Scale (WAIS) had strong validity ( $r \cong .57$ ) for predicting a range of criterion measures. In a meta-analysis of 67 studies, Chouinard and Braun (1993) documented the utility of brief neuropsychological tests to screen clients for a wide range of brain dysfunction.

### *Self-Report Measures*

Self-report tests also describe current levels of symptomatology and aid in differential diagnosis. Three Minnesota Multiphasic Personality Inventory (MMPI) meta-analyses have examined the validity of the MMPI as a descriptive device. The first two (Atkinson, 1986; Parker et al., 1988) found that the MMPI had strong validity (mean  $r \cong .42$ ) as a descriptor of personality. The third (Zalewski & Gottesman, 1991) was a comprehensive review of the ability of the MMPI to help with differential diagnosis (i.e., neurotic vs. psychotic disorders, depression vs. anxiety disorder, schizophrenia vs. affective disorder, and nonpatient vs. psychiatric patient). The authors pooled data from 403 patient and control samples (incorporating data from more than 20,000 participants) and found that the MMPI could effectively differentiate among various disorders when a configural approach to the MMPI data was used (i.e., when multiple scales were considered simultaneously). Similar findings have been reported for the revised test, the MMPI-2 (Ben-Porath, Butcher, & Graham, 1991).

Ganellen (1996) reviewed and synthesized research to clarify the diagnostic efficiency of the MMPI and Millon Clinical Multiaxial Inventory-II (MCMI-II). Ganellen calculated a number of variables, although the most important refer to (a) the probability that a diagnosable condition is present when the test scores indicate it is present (i.e., positive predictive power) and (b) the probability that a condition is absent when the test scores indicate it is absent (i.e., negative predictive power). Ganellen found that with respect to diagnosing depression, both the MMPI and MCMI-II had good positive predictive power and good negative predictive power.

Results were slightly different when considering the diagnosis of schizophrenia. Both tests had strong values for negative predictive power, but more moderate positive predictive values were found, particularly for the MMPI. Thus the tests accurately indicated when schizophrenia was absent, even though conditions other than schizophrenia (e.g., psychotic depression) also produced a psychotic-appearing MMPI or MCMI-II profile. Many other studies have provided evidence on the ability of self-report scales to differentiate various conditions on Axis I (clinical disorders) or Axis II (personality disorders) of the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; *DSM-IV*; American Psychiatric Association, 1994; e.g., Gartner, Hurt, & Gartner, 1989; Teitelbaum & Carey, 1996; Wetzler & Marlowe, 1993).

The validity of self-report tests of personality and mood has also been demonstrated in medical contexts. In a meta-analytic review, Herbert and Cohen (1993a) demonstrated that tests of depression were associated with immunological parameters such as decreased natural killer-cell activity and lymphocyte proliferation ( $r \cong .10-.25$ ). Meta-analysis has also demonstrated that self-report tests of stress predict reduced natural killer-cell activity ( $r \cong .23$ ; Herbert & Cohen, 1993b), and replicated research has demonstrated similar immunological correlates for tests of repressive coping styles (O'Leary, 1990). Finally, in an important large-scale study ( $N = 11,242$ ), Wells et al. (1989) demonstrated that self-reported depression was associated with a host of physical and emotional indicators of dysfunction. Patients with assessed depression were more functionally impaired than patients with a range of chronic medical problems (e.g., hypertension, diabetes, arthritis, angina, back problems).

### *Performance Measures of Personality*

Several meta-analytic reviews have demonstrated the utility of the Rorschach to describe symptomatology (Atkinson, 1986; Atkinson, Quarrington, Alp, & Cyr, 1986; Parker, 1983; Parker et al., 1988). In aggregate, these meta-analyses have found that when the Rorschach is used in research conducted with a sound rationale, it produces large validity coefficients (i.e.,  $r \cong .37$ ). It is important to note that the validity coefficients produced by the Rorschach are as large as those found for objectively scored tests like the MMPI. Narrative reviews on the utility of the Rorschach or Thematic Apperception Test (TAT) for descriptive and diagnostic purposes can be found in Stricker and Healey (1990), McClelland (1989), Dies (1994), Bornstein (1996), Cramer (1996), or Weiner (1994, 1996). A meta-analysis documenting the validity of the TAT for predicting functional or adaptive behaviors in a nonclinical setting can be found in Spengler (1992).

Other, more general studies have demonstrated the ability of the Rorschach or the TAT (a) to differentiate among Axis II conditions like borderline, antisocial, narcissistic, and schizotypal personality disorders and Axis I conditions like schizophrenia, major depression, conduct disorder, and panic disorder and (b) to also identify non-*DSM* conditions, such as differentiating patients who have experienced physical or sexual trauma from those who have not (e.g., de Ruiter & Cohen, 1992; Exner, 1986; Gacono & Meloy, 1994; Gartner et al., 1989; Hilsenroth, Fowler, & Padawer, 1998; Hilsenroth, Fowler, Padawer, & Handler, 1997; Leavitt & Labott, 1997; Leifer, Shapiro, Martone, & Kassem, 1991; Nigg, Lohr, Westen, Gold, & Silk, 1992). For example, Hilsenroth et al. (1998) found that the Rorschach Schizophrenia Index had excellent positive predictive power for diagnosing psychotic disorders and moderate-to-good negative predictive power.

Performance-based personality instruments also can describe clinical symptomatology. In two replication studies, one conducted with adults (Urist, 1977) and the other with adolescents (Urist &

<sup>3</sup> This coefficient was calculated from a Cohen's  $d$  value of 1.87. It assumes that there was an equal number of participants in the control and patient conditions. Typically, the coefficient would be smaller in practice because most clinical settings do not contain an equal number of patients with and without dementia. If the incidence of dementia in a clinical setting were actually 10% rather than 50%,  $r$  would be .49.

Shill, 1982), a Rorschach measure of problematic interpersonal relationships was strongly associated ( $r \cong .45-.50$ ) with clinician ratings of problematic interpersonal behavior observed on an inpatient ward. Perry, Moore, and Braff (1995) found a strong relationship ( $r = .42$ ) between ego deficits as measured by the Rorschach and general impairment in social and occupational functioning, even though clinician ratings of symptomatology were not similarly effective at predicting this criterion. Other researchers have found similar positive associations using Rorschach or TAT measures with children (e.g., Cramer, 1996; Tuber, 1992). Also, Burns and Viglione (1996) found a Rorschach measure of interpersonal relatedness had a strong ability to predict social functioning in a nonclinical sample. This measure predicted functioning over and above other variables such as education, intelligence, age, income, and use of psychiatric medication.

### Description and Prediction of Functional Behavior

Functional behavior can include a patient's capacity to work effectively, parent effectively, work without missing days due to illness, and so on. The capacity of a patient to engage in functional behaviors can have an effect on the patient's diagnosis, treatment, and prognosis. Psychological assessments can describe and quantify functional behavior for diagnostic, treatment, and prognostic purposes and can extend to behaviors that are relevant outside an acute medical or psychiatric context. Functional behavior is receiving increased emphasis in clinical and other settings, in part because of the increasing participation of third and fourth parties in medical and psychiatric treatment (Moras, 1997). The practical everyday applications of psychological assessment to the rehabilitation process, to vocational and educational planning, to determinations of legal competence, and to making determinations regarding appropriate living arrangements for people who are cognitively or emotionally impaired represent additional ways in which psychological assessment can be useful in describing and predicting functional behavior (e.g., Loeb, 1996; Sbordone & Long, 1996).

### Tests of Cognitive Ability

Extensive meta-analytic research, synthesizing data from thousands of studies and hundreds of thousands of participants, has documented that tests of cognitive ability are highly predictive of proficiency on the job and success in job training, particularly when job functioning requires complex skills (uncorrected  $r \cong .35-.45$ ; see, e.g., Gottfredson, 1997; Hunter, 1986; Hunter & Hunter, 1984; Kaufman, 1990; Schmitt, Gooding, Noe, & Kirsch, 1984). Extensive research on tests of cognitive ability has demonstrated their strong utility as descriptors and predictors of academic achievement ( $r \cong .45-.60$ ; for up-to-date reviews, see Gottfredson, 1997, or Kaufman, 1990).

Other review articles and primary studies have indicated that scores on cognitive and neuropsychological tests have moderate-to-strong relationships with everyday functional behaviors such as self-help skills, independent living skills, driving ability, academic success, vocational skills, and employment status (e.g., Crepeau & Scherzer, 1993; Gottfredson, 1997; Lezak, 1995; Olea & Ree, 1994; Ree, Earles, & Teachout, 1994; Sbordone & Long, 1996; Verive & McDaniel, 1996). For example, in a cross-validated

study, Heaton, Chelune, and Lehman (1978) found that the Halstead-Reitan Neuropsychological Battery (HRNB), which includes the WAIS and the MMPI, was able to clearly differentiate patients who had the functional capacity to work from those who did not (correctly classifying 81% of the unemployed patients and 86% of the patients employed full time). A longitudinal study using a similar design found that the same tests were highly correlated with employment status, income, and job skills 6 months after the baseline evaluation (Newnan, Heaton, & Lehman, 1978).

### Self-Report Personality Measures

Meta-analytic reviews incorporating data from tens of thousands of participants have documented the validity and utility of self-report personality tests as indicators of functional ability and performance on daily tasks (e.g., Barrick & Mount, 1991; Hough, 1992; Hunter & Hunter, 1984; Schmitt et al., 1984; Tett, Jackson, & Rothstein, 1991). Although these associations are generally in the small-to-moderate range (uncorrected  $r \cong .10-.20$ ), such studies unequivocally demonstrate that personality test scores predict behaviors such as absenteeism from work, success on the job or in training, problem-solving ability, creativity, disciplinary problems, irresponsible behavior, initiative, and so on. It is important to note that, in both nonclinical settings (Robertson & Kinder, 1993) and medical or psychiatric settings (Heaton et al., 1978; Newnan et al., 1978), personality testing contributes useful information to the prediction of these criterion behaviors that is independent of cognitive ability. In other words, it is often best to use both personality and cognitive tests to predict these behaviors, suggesting that multimethod assessment batteries are more valuable than single-method approaches to assessment.

### Independent Living Scales and Situational Tests

Recent efforts have focused on developing psychological tests that are specifically designed to predict functional behaviors away from a medical or psychiatric context. For example, the Independent Living Scales (ILS; Loeb, 1996) test is a performance-based assessment tool for evaluating adults' competence to engage in independent daily living activities. Its content was developed by subject-matter experts who identified the critical behaviors people need in taking care of themselves and their property. Validation research to date indicates that the scales can be used to identify strengths and weaknesses in carrying out activities of daily living and to design treatment interventions or make decisions regarding appropriate living arrangements.

Other measures include the Rivermead Behavioral Memory Test (Wilson, Cockburn, Baddeley, & Hiorns, 1989) and the Behavioral Inattention Test (Wilson, Cockburn, & Halligan, 1987), both of which evaluate patients by having them complete tasks in an office setting that are analogous to everyday, real-world tasks (e.g., recalling a new route, delivering a message, telephone dialing, menu reading, coin sorting). These measures have demonstrated impressive strength for predicting functional behaviors outside of the clinic (see Lezak, 1995, for additional measures and data).

## Prediction of Health Outcomes

Psychological assessments have demonstrated utility in predicting a wide range of health and medical outcomes. Furthermore, such utility has been demonstrated for a wide range of measures.

### *Self-Report Personality Tests*

*General health outcomes.* Several meta-analyses have examined the ability of psychological tests to predict subsequent health-related outcomes. In a review of 56 studies, Holden (1991) determined that baseline assessment of a patient's perceived self-efficacy predicted outcomes related to smoking cessation, pain tolerance, weight loss, adherence to exercise regimes, and tolerance of dental procedures ( $r \cong .26$ ). Booth-Kewley and Friedman (1987) found that baseline assessment of depression, anger, or anxiety predicted subsequent heart-disease outcomes (i.e., angina, atherosclerosis, myocardial infarction, cardiac death). Although the magnitude of these relationships was small ( $r \cong .15$ ), these psychological variables were as important to the prediction of outcome as traditional risk factors such as cholesterol or smoking (Friedman & Booth-Kewley, 1987). In a broader meta-analytic review, Friedman and Booth-Kewley (1987) extended their findings by documenting the relationship between these personality variables and other disease entities such as asthma, ulcer, arthritis, and headache ( $r \cong .10-.20$ ). S. Cohen and Williamson (1991) conducted a review of factors associated with susceptibility to infectious disease. They identified a number of studies indicating that the baseline assessment of negative emotional states or social introversion could be used to predict subsequent utilization of medical services, subsequent complaints of illness, slowed recovery from upper respiratory infection, susceptibility to viral infection, recurrence of herpes, severity of response to bacterial infection, and decreased immune functioning.

*Organ transplantation.* In organ transplant cases, baseline personality testing has been shown to predict survival, treatment adherence, and health care utilization (Chacko, Harper, Gotto, & Young, 1996; Shapiro et al., 1995). For example, Chacko et al. (1996) determined that pretreatment psychological testing was a better predictor of transplant survival than any interview-based measures. In fact, on the basis of pretreatment testing, high-risk and low-risk groups could readily be identified. Over a follow-up period of 9 to 56 months, the low-risk group had a cumulative survival rate of approximately 88%, whereas the high-risk group had a cumulative survival rate of only 48%.

*Asthma.* Dirks and his colleagues (Dirks, 1982; Dirks & Kinsman, 1981; Dirks, Kinsman, Horton, Fross, & Jones, 1978; Dirks et al., 1977; Dirks, Robinson, & Moore, 1981; Dirks, Schraa, Brown, & Kinsman, 1980) carried out an extensive series of studies that examined the utility of the MMPI for predicting asthmatic outcomes. Specifically, they identified an MMPI measure of characterological anxiety and ego resources that is unrelated to the physical severity of asthma, yet effectively predicts the mislabeling of bodily sensations as asthma attacks, the extent to which patients are inaccurately seen as impaired by their physicians, the amount of unnecessary steroid medication prescribed at hospital discharge, compliance with medication following discharge, length of hospitalization, the frequency of rehospitalization, and the excess cost of rehospitalization (e.g., Dirks et al.,

1978; Dirks et al., 1981; Dirks et al., 1980). For example, Dirks et al. (1980) demonstrated that over a 2-year follow-up period, the patients that were predicted to have worse outcomes were rehospitalized for nearly 3 times as many days as the patients who were predicted to have better outcomes (an average of 31.3 days vs. 11.1 days). It is important to note that rehospitalization rates were not associated with medical parameters of illness severity (Dirks et al., 1980), and longitudinal outcomes could not be predicted from baseline measures of pulmonary function, physician judgments of the severity of medical problems, or the use of oral corticosteroids at initial hospital discharge (Dirks et al., 1978).

As would be expected, the ability of this MMPI measure to predict outcome was enhanced when it was combined with other sources of information so that clinical judgments were derived from a battery of test results (e.g., Dirks, 1982; Dirks & Kinsman, 1981). Furthermore, this MMPI measure of ego resources has demonstrated its usefulness for predicting outcome in other medical conditions, including obesity and forms of respiratory illness other than asthma (e.g., Dirks et al., 1977; Flanagan & Wagner, 1991).

### *Performance-Based Personality Tests*

Research has also documented the value of performance-based personality assessment methods for predicting health and medically related outcomes. McClelland (1989) reviewed the literature documenting the association between variables assessed by the TAT and immune functioning or illness (also see McKay, 1991; and O'Leary, 1990). Constellations of variables measured by the TAT have been associated ( $r \cong .25-.50$ ) with the development of high blood pressure; increased frequency of minor illnesses, particularly upper respiratory infections; and increased frequency of major illnesses. Although the mechanisms responsible for these findings have not been completely determined, they are thought to be a function of the associations that have been found ( $r \cong .20-.50$ ) between various TAT variables and norepinephrine release, salivary immunoglobulin A levels, natural killer-cell activity levels, and the ability or inability to reduce physical symptoms through psychological treatment.

The Precursors Study carried out by Thomas and her colleagues (e.g., Graves, Phil, Mead, & Pearson, 1986; Thomas, 1988) has provided relevant data regarding the Rorschach. In this extensive longitudinal study, 1,154 healthy students were tested at entrance into medical school and were followed for 18 to 35 years. Though the magnitudes of the associations were small ( $r \cong .05-.20$ ), baseline Rorschach scores predicted subsequent cancer, mental illness, suicide, physical ill health, and all-cause mortality. At least with respect to cancer, Rorschach data provided incremental information that could not be obtained from other sources, predicting cancer occurrence over and above other risk factors such as smoking history, serum cholesterol level, and time elapsed since baseline testing (Graves et al., 1986). After correcting for these potentially confounding factors, the physicians who were predicted to have the highest incidence of cancer were 3 to 4 times more likely to develop cancer than the physicians who were predicted to have the lowest incidence.

### *Neuropsychological and Cognitive Tests*

Neuropsychological tests have also demonstrated utility as predictors of health and medical outcome. For instance, neuropsychological tests have predicted outcome for the surgical treatment of epilepsy. This surgery is costly, involves risk, and for a sizable number of patients does not result in relief from seizures. Neuropsychological assessment has been helpful in identifying those who would or would not be helped by this procedure (Chelune, Naugle, Luders, & Awad, 1991; Dodrill, Wilkus, & Ojemann, 1992). In addition, several studies have documented that baseline neuropsychological testing can usefully predict the subsequent onset of dementia in otherwise nonsymptomatic patients (e.g., Crystal et al., 1996) or in patients who had age-associated memory impairment on initial evaluation (e.g., Hanninen et al., 1995).

### Prediction of Health Care Utilization

Several individual studies demonstrate that psychological assessment instruments can identify patients who are likely to utilize health care services more often than average. These studies are pertinent because health care utilization clearly influences third-party payer "bottom line" decision making.

In a well-designed study, Bornstein, Krukoni, Manning, Mastrosimone, and Rossner (1993) found that the assessment of dependency in college students was a strong predictor of subsequent health care utilization. Over the course of 3 months, relative to those who had low baseline dependency scores, those high in dependency made 2.6 times as many visits to a campus health center and 4.6 times as many visits to a private physician. It is important to note that these powerful predictive relationships held even after controlling for overall physical health. Similar results were obtained in a study examining the length of stay for general medical patients in a tertiary-care facility (Saravay, Steinberg, Weinschel, Pollack, & Alovio, 1991). The baseline assessment of negative emotionality and cognitive functioning predicted subsequent length of stay in the hospital even after controlling for each patient's extent of physical impairment. It is significant that these psychological variables were much better predictors than such variables as medical diagnosis, occupation, preadmission living arrangements, and so on.

Tessler, Mechanic, and Dimond (1976) conducted a long-term study of physician utilization in a prepaid medical plan. Baseline testing indicated that psychological distress predicted the number of visits made to a physician over the subsequent year even after controlling for a variety of physical health variables. Allison et al. (1995) demonstrated that the baseline assessment of psychological distress in a sample of cardiac rehabilitation patients predicted subsequent rehospitalization rates, as well as subsequent cardiac death, myocardial infarction, or cardiac arrest with resuscitation. As would be expected, baseline testing also predicted the subsequent costs for rehospitalization. On average, those who were emotionally distressed at baseline had subsequent 6-month hospital charges of \$9,504, compared with \$2,146 in charges for those who were not distressed at baseline testing. Finally, in at least one study, the baseline assessment of cognitive ability has predicted subsequent use of inpatient and outpatient health care resources (LoGiudice et al., 1997). Over the course of 1 year, those patients with impaired cognitive processes at baseline had total health care

costs that were twice as high as the patients without cognitive difficulties.

### Prediction of Psychotherapy, Forensic, and Mental Health Outcomes

Predicting psychotherapy, forensic, and general mental health and behavioral outcomes has proven to be a formidable task (Garfield, 1994; Luborsky, Chandler, Auerbach, Cohen, & Bachrach, 1971). Nevertheless, evidence exists for the efficacy of personality tests for these purposes.

### *Psychotherapy Outcomes*

One of the best predictors of psychotherapy outcome comes from the Rorschach Prognostic Rating Scale (RPRS), which was designed to predict which patients would be successful in psychotherapy. Meyer and Handler (1997) recently performed a meta-analysis on the existing literature that used the RPRS to predict subsequent outcome. They found that the RPRS had a powerful ability ( $r = .44$ ) to predict psychotherapy outcomes approximately 1 year after baseline testing. In fact, the ability of this Rorschach scale to predict outcome was much stronger than the ability of many other medical, psychological, or educational tests to do so. For example, the RPRS demonstrated better predictive validity than the Cardiac Stress Test for predicting subsequent cardiac disease, blood pressure testing for predicting the recurrence of a heart attack, or SAT scores for predicting college success, to name a few.

In two studies of children, Tuber (1983) and LaBarbera and Cornsweet (1985) found that Rorschach scores were strong predictors of outcome and rehospitalization rates ( $r = .45$ ). In adults, Rorschach scores have predicted ratings of the ability to engage in short-term dynamic therapy (Alpher, Henry, & Strupp, 1990) as well as the outcome from this form of treatment (Alpher, Perfetto, Henry, & Strupp, 1990). Rorschach scores of ego strength have also demonstrated a powerful ability to predict response to antidepressant treatment. In both a 9-week ( $r = .57$ ) and a 5-year ( $r = .42$ ) follow-up study (Perry, McDougall, & Viglione, 1995; Perry & Viglione, 1991), measures of ego strength could predict outcome over and above such other variables as age, gender, education, baseline severity of depression, and frequency of treatment visits. Greenberg and Bornstein (1989) have also documented that a Rorschach scale of dependency predicts length of psychiatric hospitalization. This relationship was particularly strong for women; those high on the scale had a mean hospital stay that was nearly twice as long as women who were low on the scale.

Recent adult research has indicated that pretreatment testing with self-report scales (Kopta, Howard, Lowry, & Beutler, 1994) or the Rorschach (Exner & Andronikof-Sanglade, 1992; Weiner & Exner, 1991) can differentiate symptoms that change quickly in treatment from those that require treatment of moderate duration, and both of these can be differentiated from the characterological symptoms that respond only to lengthy treatment. Finally, a recent study indicated that baseline MCMI-II scores could predict psychiatric rehospitalization rates over a 2-year follow-up period, even though other variables, such as *DSM-IV* Axis I (American Psychiatric Association, 1994) diagnoses, could not (Donat, 1997).

### Forensic Outcomes

A meta-analytic review offered support for the ability of the Hare Psychopathy Checklist (PCL) to predict outcomes in a criminal context (Salekin, Rogers, & Sewell, 1996). The PCL was able to predict subsequent violent behavior ( $r = .33$ ), recidivism ( $r = .28$ ), and sexual sadism or deviant sexual arousal ( $r = .25$ ).<sup>4</sup> In addition, the PCL has proved to be a better predictor of these outcomes than virtually all other variables or procedures, including criminal history, psychiatric history, age, substance use, or a diagnosis of antisocial personality disorder (Bonta, Law, & Hanson, 1998; Hanson & Bussire, 1998), making it a sound and useful tool for assessing these very hard-to-predict outcomes.

### Mental Health Outcomes

Consistent or promising results in relating psychological tests to mental health and behavioral outcomes have been observed in a number of individual studies using a variety of assessment methods. For example, in both children and adults, baseline testing of self-reported personality has indicated that manifestations of the broad neuroticism–negative emotionality trait predict negative outcomes. Studying children in a residential setting for a period of up to 4 years, Mattison and colleagues (e.g., Mattison, Handford, Kales, Goodman, & McLaughlin, 1990) found that baseline self-reports of negative emotionality predict subsequent disciplinary action, classroom failure, reports of behavior problems by teachers and house-parents, greater need for counseling, and more premature departures from treatment. In adults, baseline testing has determined that self-reported neuroticism is a better predictor of long-term clinical outcome in depression than a host of other potential predictors, including age, social or marital status, prior depressions, prior suicide attempts, pretreatment depression levels, and pretreatment stress levels (Hirschfeld, Klerman, Andreasen, Clayton, & Keller, 1986; Weissman, Prusoff, & Klerman, 1978). In general, elevated baseline neuroticism scores predispose people to experience a range of negative life outcomes (Magnus, Diener, Fujita, & Pavot, 1993) as well as negative outcomes in individual and marital therapy (Luborsky et al., 1993; Mohr et al., 1990; Snyder, Mangrum, & Wills, 1993).

Studying characteristics other than neuroticism, Achenbach and colleagues (e.g., Achenbach, Howell, McConaughy, & Stanger, 1995a, 1995b) conducted a large-scale study on a nationally representative sample of children ( $N = 2,466$ ). They relied on self-report tests and parent or teacher rating scales to predict outcomes that were measured up to 6 years later. Baseline test data predicted (with cross-method  $r$ s ranging from .10 to .25) subsequent academic problems, school behavior problems, use of mental health services, suicidal thinking, police contacts, and problematic behavioral syndromes (e.g., withdrawal, social problems, attention problems). In adults, baseline testing data obtained from observer ratings have been found to predict rehospitalization rates as well as successful return to work (e.g., Prigatano et al., 1984; Sappington & Michaux, 1975). Replicated research has also revealed that baseline assessment of hopelessness is the best predictor of subsequent suicide (e.g., Beck, Brown, Berchick, Stewart, & Steer, 1990). Finally, Bornstein, Bowers, and Robinson (1995) found that, as hypothesized, a Rorschach scale of dependency was a very strong predictor ( $r = .84$ ) of significant interpersonal life events (e.g., changes in relationships) over a subsequent 3-month period.

### Identification of Patient Characteristics That Affect Treatment

Psychological treatment can vary across a variety of dimensions including theoretical framework (psychodynamic to ecological–multisystemic), duration (brief to long term), focus (symptom management to personality restructuring), and therapist stance (nondirective to directive). To the extent that psychological assessment instruments identify patient characteristics that can affect the success of different types of psychological treatment, they can be used to enhance treatment outcome by facilitating the match of patient characteristics to treatment types.

Carefully conceptualized and implemented studies have indicated that patients with externalizing symptoms (acting out, projecting, avoidance defenses) do better in treatment that is more structured or directed by the therapist, while patients with internalizing symptoms (self-punishment, anxiousness, worry) do better in treatments where they set the pace and determine the structure (e.g., Beutler, Machado, Engle, & Mohr, 1993; Blatt, 1992).

Research examining the match between patient characteristics and forms of treatment is still at a relatively early stage of development, characterized by a diverse array of assessment methods, measured constructs, treatments, and outcome variables. Although results are not uniformly positive, a number of investigators have documented the value of pretreatment test data for predicting treatment outcome (e.g., Crits-Christoph, Barber, & Kurcias, 1993; DuPaul & Ervin, 1996; Eifert, Evans, & McKendrick, 1990; Project MATCH Research Group, 1997; Schulte, 1996). In separate reviews of the literature, Garfield (1994) and Piper (1994) concluded that psychological testing probably has the greatest predictive value for individual or group psychotherapy when oriented toward identifying the best match between patient characteristics and therapist–treatment characteristics. Dies and Dies (1995) articulated a model for incorporating testing into the various stages of group treatment in order to improve service delivery, and Haynes, Leisen, and Blaine (1997) reviewed the literature (see their Table 2) and articulated models for designing behavioral interventions based on functional assessments. Overall, these data and reviews speak to the importance of individualized assessment for matching patients to optimal treatment.

### Using Psychological Assessment as a Treatment in Itself

Another use of psychological assessment is as a brief treatment in itself. Besides being used to plan, monitor, and evaluate clinical interventions, assessment can be a “treatment in microcosm” with a positive effect on patients (Allen, 1981, p. 251). In fact, many case reports indicate that psychological testing can have a dramatic effect on patients’ symptomatology and on their understanding of themselves and their problems (e.g., Clair & Prendergast, 1994; Finn, 1996; Finn & Martin, 1997; Fischer, 1985/1994; Moffett, Steinberg, & Rohde, 1996).

Recently, two controlled studies have documented the effects of psychological assessment as a brief treatment. Finn and Tonsager (1992) found that patients receiving a 2-hour MMPI-2 assessment

<sup>4</sup> All values reported here were computed from original data that predicted subsequent behavior. Coefficients were weighted by sample size.

with feedback reported less symptomatology, higher self-esteem, and greater hopefulness about their problems compared with a control group. Patients in the control group also completed the MMPI but they only talked to the therapist about their current concerns and did not receive test feedback. Symptomatic improvement in the group receiving assessment and feedback was observed both immediately following the assessment and at a 2-week follow-up. The effects found in this study were not only statistically significant but also clinically meaningful (at follow-up, mean  $r = .35$ ). Newman and Greenway (1997) replicated and extended the study by Finn and Tonsager (1992) in a sample of Australian outpatients, finding similar improvements in self-esteem and symptomatology (at follow-up, mean  $r = .27$ ). Hanson, Claiborne, and Kerr (1997) showed that patients invited to verify and elaborate feedback from personality and vocational tests considered their interpretive sessions to be deeper and their assessors to be more expert, attractive, and trustworthy than did patients who were given feedback in a noncollaborative manner.

Case studies suggest that psychological assessment is an effective intervention not only with individual patients but also with couples, families, and other interpersonal systems. For example, Dorr (1981) detailed a brief marital therapy centered on psychological assessment. He concluded that with a focal assessment it "often is possible to accomplish in 1½ to 2½ days what would require weeks or months of once-a-week sessions" (p. 554). Fulmer, Cohen, and Monaco (1985); Quirk, Storsahl, Kreilkamp, and Erdberg (1995); and Pollak (1988) presented case examples where psychological testing was used to help parents and school personnel gain empathy for children with learning and behavioral problems and to develop effective interventions. Moffett et al. (1996) described the use of personality testing to assist the formation of a "therapeutic community" among men in inpatient treatment for severe substance dependence. Berg (1988) and L. J. Cohen (1980) discussed how testing could be used to analyze and resolve interpersonal difficulties that arise between difficult patients and their therapists or treatment teams.

### Conclusion

The numerous studies that we have reviewed represent a sample of the extensive research base that supports the validity and utility of psychological assessment instruments for a range of applications in clinical health care today. Our full PAWG report (Meyer et al., 1998) and a companion article (Meyer et al., 1999) provide even more empirical support for the validity of psychological tests.

Yet, third-party payers continue to deny and minimize the validity and utility of psychological assessment in clinical health care decision making. In today's competitive health care marketplace, psychological assessment has been maligned as costly and ineffective, with third-party payer authorization and appropriate reimbursement increasingly difficult to obtain.

We have concluded that we, as psychologists, must share some responsibility for this problem. In general, we have failed to educate skeptical third-party payers about the considerable empirical support that exists for the validity of psychological assessment instruments. We may have failed to do so, at least in part, because of the voluminous, highly technical nature of the large body of scientific literature that has grown around psychological assessment over many decades. The breadth and technical complexity of

this research has made it difficult for practitioners to access and clearly convey to decision makers the extent of the empirical support that psychological assessment instruments enjoy.

We hope that this article, in conjunction with the additional data reported by Meyer et al. (1998, 1999), helps us toward a solution to this vexing problem. By consolidating in understandable terms some of this broad research area we hope to provide practitioners with accessible, useful information that can be used to meaningfully, consistently, and accurately communicate with third-party decision makers when psychological assessment authorization and reimbursement is being considered. To the extent that third-party payers become convinced that psychological assessment can contribute to cost-effective health care, they are likely to support its appropriate use.

We are not so naive as to suggest that the road to greater acceptance and valuing of psychological assessment is a smooth one, nor that the road leads straight to greater support for psychological assessment. Nevertheless, over time, we believe that a sustained grassroots effort to educate third-party payers by informed, assessment-minded psychologists, coupled with ongoing advocacy initiatives by state and national psychological associations, can arrest and even reverse the recent decline of the practice of psychological assessment.

### References

- Achenbach, T. M., Howell, C. T., McConaughy, S. H., & Stanger, C. (1995a). Six-year predictors of problems in a national sample of children and youth: II. Signs of disturbance. *Journal of the American Academy of Child and Adolescent Psychiatry, 34*, 488-498.
- Achenbach, T. M., Howell, C. T., McConaughy, S. H., & Stanger, C. (1995b). Six-year predictors of problems in a national sample of children and youth: III. Transitions to young adult syndromes. *Journal of the American Academy of Child and Adolescent Psychiatry, 34*, 658-669.
- Allen, J. G. (1981). The clinical psychologist as diagnostic consultant. *Bulletin of the Menninger Clinic, 45*, 247-258.
- Allison, T. G., Williams, D. E., Miller, T. D., Patten, C. A., Bailey, K. R., Squires, R. W., & Gau, G. T. (1995). Medical and economic costs of psychologic distress in patients with coronary artery disease. *Mayo Clinic Proceedings, 70*, 734-742.
- Alpher, V. S., Henry, W. P., & Strupp, H. H. (1990). Dynamic factors in patient assessment and prediction of change in short-term dynamic psychotherapy. *Psychotherapy, 27*, 350-361.
- Alpher, V. S., Perfetto, G. A., Henry, W. P., & Strupp, H. H. (1990). The relationship between the Rorschach and assessment of the capacity to engage in short-term dynamic psychotherapy. *Psychotherapy, 27*, 224-229.
- Ambrose, P. A., Jr. (1997). Challenges for mental health service providers: The perspective of managed care organizations. In J. N. Butcher (Ed.), *Personality assessment in managed health care* (pp. 61-72). New York: Oxford University Press.
- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- Atkinson, L. (1986). The comparative validities of the Rorschach and MMPI: A meta-analysis. *Canadian Psychology, 27*, 238-247.
- Atkinson, L., Quarrington, B., Alp, I. E., & Cyr, J. I. (1986). Rorschach validity: An empirical approach to the literature. *Journal of Clinical Psychology, 42*, 360-362.
- Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44*, 1-26.
- Beck, A. T., Brown, G., Berchick, R. J., Stewart, B. L., & Steer, R. A.



- (1990). Relationship between hopelessness and ultimate suicide: A replication with psychiatric outpatients. *American Journal of Psychiatry*, 147, 190–195.
- Ben-Porath, Y. S., Butcher, J. N., & Graham, J. R. (1991). Contribution of the MMPI-2 content scales to the differential diagnosis of schizophrenia and major depression. *Psychological Assessment*, 3, 634–640.
- Berg, M. (1988). Diagnostic benefits of analyzing the group dynamics of the assessment team. *Bulletin of the Menninger Clinic*, 52, 126–133.
- Beutler, L. E., Machado, P. P. P., Engle, D., & Mohr, D. (1993). Differential Patient  $\times$  Treatment Maintenance among cognitive, experiential, and self-directed psychotherapies. *Journal of Psychotherapy Integration*, 3, 15–31.
- Blatt, S. J. (1992). The differential effect of psychotherapy and psychoanalysis with anaclitic and introjective patients: The Menninger Psychotherapy Research Project revisited. *Journal of the American Psychoanalytic Association*, 40, 691–724.
- Bonta, J., Law, M., & Hanson, K. (1998). The prediction of criminal and violent recidivism among mentally disordered offenders: A meta-analysis. *Psychological Bulletin*, 123, 123–142.
- Booth-Kewley, S., & Friedman, H. S. (1987). Psychological predictors of heart disease: A quantitative review. *Psychological Bulletin*, 101, 343–362.
- Bornstein, R. F. (1996). Construct validity of the Rorschach Oral Dependency Scale: 1967–1995. *Psychological Assessment*, 8, 200–205.
- Bornstein, R. F., Bowers, K. S., & Robinson, K. J. (1995). Differential relationships of objective and projective dependency scores to self-reports of interpersonal life events in college student subjects. *Journal of Personality Assessment*, 65, 255–269.
- Bornstein, R. F., Krukonis, A. B., Manning, K. A., Mastrosimone, C. C., & Rossner, S. C. (1993). Interpersonal dependency and health service utilization in a college student sample. *Journal of Social and Clinical Psychology*, 12, 262–279.
- Burns, B., & Viglione, D. J., Jr. (1996). The Rorschach Human Experience Variable, interpersonal relatedness, and object representation in nonpatients. *Psychological Assessment*, 8, 92–99.
- Chacko, R. C., Harper, R. G., Gotto, J., & Young, J. (1996). Psychiatric interview and psychometric predictors of cardiac transplant survival. *American Journal of Psychiatry*, 153, 1607–1612.
- Chelune, G. J., Naugle, R. I., Luders, H., & Awad, I. A. (1991). Prediction of cognitive change as a function of preoperative ability status among temporal lobectomy patients seen at 6-month follow-up. *Neurology*, 41, 399–404.
- Chouinard, M.-J., & Braun, C. M. (1993). A meta-analysis of the relative sensitivity of neuropsychological screening tests. *Journal of Clinical and Experimental Neuropsychology*, 15, 591–607.
- Christensen, D., Hadzi-Pavlovic, D., & Jacomb, P. (1991). The psychometric differentiation of dementia from normal aging: A meta-analysis. *Psychological Assessment*, 3, 147–155.
- Clair, D., & Prendergast, D. (1994). Brief psychotherapy and psychological assessments: Entering a relationship, establishing a focus, and providing feedback. *Professional Psychology: Research and Practice*, 25, 46–49.
- Cohen, L. J. (1980). The unstated problem in a psychological testing referral. *American Journal of Psychiatry*, 137, 1173–1176.
- Cohen, S., & Williamson, G. M. (1991). Stress and infectious disease in humans. *Psychological Bulletin*, 109, 5–24.
- Cramer, P. (1996). *Storytelling, narrative, and the Thematic Apperception Test*. New York: Guilford Press.
- Crepeau, F., & Scherzer, P. (1993). Predictors and indicators of work status after traumatic brain injury: A meta-analysis. *Neuropsychological Rehabilitation*, 3, 5–35.
- Crits-Christoph, P., Barber, J. P., & Kurcias, J. S. (1993). The accuracy of therapists' interpretations and the development of the therapeutic alliance. *Psychotherapy Research*, 3, 25–35.
- Crystal, H. A., Dickson, D., Sliwinski, M., Masur, D., Blau, A., & Lipton, R. B. (1996). Associations of status change measures of neuropsychological function with pathological changes in elderly, originally nondemented subjects. *Archives of Neurology*, 53, 82–87.
- de Ruiter, C., & Cohen, L. (1992). Personality in panic disorder with agoraphobia: A Rorschach study. *Journal of Personality Assessment*, 59, 304–316.
- Dies, R. R. (1994). The Rorschach Comprehensive System: Current status and clinical applications. *Independent Practitioner*, 14, 96–106.
- Dies, R. R., & Dies, K. R. (1995). The role of evaluation in clinical practice: Overview and group treatment illustration. *International Journal of Group Psychotherapy*, 43, 77–105.
- Dirks, J. F. (1982). Bayesian prediction of psychomaintenance related to rehospitalization in asthma. *Journal of Personality Assessment*, 46, 159–163.
- Dirks, J. F., & Kinsman, R. A. (1981). Clinical prediction of medical rehospitalization: Psychological assessment with the Battery of Asthma Illness Behavior. *Journal of Personality Assessment*, 45, 608–613.
- Dirks, J. F., Kinsman, R. A., Horton, D. J., Fross, K. H., & Jones, N. F. (1978). Panic-fear in asthma: Rehospitalization following intensive long-term treatment. *Psychosomatic Medicine*, 40, 5–13.
- Dirks, J. F., Kinsman, R. A., Jones, N. F., Spector, S. L., Davidson, P. T., & Evans, N. W. (1977). Panic-fear: A personality dimension related to length of hospitalization in respiratory illness. *Journal of Asthma Research*, 14, 61–71.
- Dirks, J. F., Robinson, S. K., & Moore, P. N. (1981). The prediction of psychomaintenance in chronic asthma. *Psychotherapy and Psychosomatics*, 36, 105–115.
- Dirks, J. F., Schraa, J. C., Brown, E. L., & Kinsman, R. A. (1980). Psycho-maintenance in asthma: Hospitalization rates and financial impact. *British Journal of Medical Psychology*, 53, 349–354.
- Dodrill, C. B., Wilkus, R. J., & Ojemann, L. M. (1992). Use of psychological and neuropsychological variables in selection of patient for epilepsy surgery. *Epilepsy Research Supplement*, 5, 71–75.
- Donat, D. C. (1997). Personality traits and psychiatric rehospitalization: A two-year follow-up. *Journal of Personality Assessment*, 68, 703–711.
- Dorr, D. (1981). Conjoint psychological testing in marriage therapy: New wine in old skins. *Professional Psychology*, 12, 549–555.
- DuPaul, G. F., & Ervin, R. A. (1996). Functional assessment of behaviors related to attention-deficit/hyperactivity disorder: Linking assessment to intervention design. *Behavior Therapy*, 27, 601–622.
- Eifert, G. H., Evans, I. M., & McKendrick, V. G. (1990). Matching treatments to client problems not diagnostic labels: A case for paradigmatic behavior therapy. *Journal of Behavior Therapy and Experimental Research*, 21, 163–172.
- Eisman, E. J., Dies, R. R., Finn, S. E., Eyde, L. D., Kay, G. G., Kubiszyn, T. W., Meyer, G. J., & Moreland, K. (1998). *Problems and limitations in the use of psychological assessment in contemporary healthcare delivery: Report of the Board of Professional Affairs Psychological Assessment Work Group, Part II*. Washington, DC: American Psychological Association.
- Eisman, E., Dies, R., Finn, S. E., Eyde, L. D., Kay, G. G., Kubiszyn, T. W., Meyer, G. J., & Moreland, K. L. (2000). Problems and limitations in the use of psychological assessment in contemporary health care delivery. *Professional Psychology: Research and Practice*, 31, 131–140.
- Exner, J. E., Jr. (1986). Some Rorschach data comparing schizophrenics with borderline and schizotypal personality disorders. *Journal of Personality Assessment*, 50, 455–471.
- Exner, J. E., Jr. (1989). Searching for projection in the Rorschach. *Journal of Personality Assessment*, 53, 520–536.
- Exner, J. E., Jr., & Andronikof-Sanglade, A. (1992). Rorschach changes following brief and short-term therapy. *Journal of Personality Assessment*, 59, 59–71.
- Finn, S. E. (1996). *A manual for using the MMPI-2 as a therapeutic intervention*. Minneapolis: University of Minnesota Press.

- Finn, S. E., & Martin, H. (1997). Therapeutic assessment with the MMPI-2 in managed healthcare. In J. N. Butcher (Ed.), *Objective psychological assessment in managed healthcare: A practitioner's guide* (pp. 131–152). New York: Oxford University Press.
- Finn, S. E., & Tonsager, S. E. (1992). The therapeutic effects of providing MMPI-2 test feedback to college students awaiting psychotherapy. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, 4, 278–287.
- Fischer, C. T. (1994). *Individualizing psychological assessment*. Hillsdale, NJ: Erlbaum.
- Flanagan, D. A., & Wagner, H. L. (1991). Expressed emotion and panic-fear in the prediction of diet treatment compliance. *British Journal of Clinical Psychology*, 30, 231–240.
- Friedman, H. S., & Booth-Kewley, S. (1987). The “disease-prone personality”: A meta-analytic view of the construct. *American Psychologist*, 42, 539–555.
- Fulmer, R. H., Cohen, S., & Monaco, G. (1985). Using psychological assessment in structural family therapy. *Journal of Learning Disabilities*, 18, 145–150.
- Gacono, C. B., & Meloy, J. R. (1994). *The Rorschach assessment of aggressive and psychopathic personalities*. Hillsdale, NJ: Erlbaum.
- Ganellen, R. J. (1996). Comparing the diagnostic efficiency of the MMPI, MCMI-II, and Rorschach: A review. *Journal of Personality Assessment*, 67, 219–243.
- Garb, H. N., & Schramke, C. J. (1996). Judgment research and neuropsychological assessment: A narrative review and meta-analyses. *Psychological Bulletin*, 120, 140–153.
- Garfield, S. L. (1994). Research on client variables in psychotherapy. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 190–228). New York: Wiley.
- Gartner, J., Hurt, S. W., & Gartner, A. (1989). Psychological test signs of borderline personality disorder: A review of the empirical literature. *Journal of Personality Assessment*, 53, 423–441.
- Gottfredson, L. S. (1997). Why g matters: The complexity of everyday life. *Intelligence*, 24, 79–132.
- Graves, P. L., Phil, M., Mead, L. A., & Pearson, T. A. (1986). The Rorschach Interaction Scale as a potential predictor of cancer. *Psychosomatic Medicine*, 48, 549–563.
- Greenberg, R. P., & Bornstein, R. F. (1989). Length of psychiatric hospitalization and oral dependency. *Journal of Personality Disorders*, 3, 199–204.
- Hanninen, T., Hallikainen, M., Koivisto, K., Helkala, E.-L., Reinikainen, K. J., Soininen, H., Mykkanen, L., Laakso, M., Pyorala, K., & Riekkinen, P. J. (1995). A follow-up study of age-associated memory impairment: Neuropsychological predictors of dementia. *Journal of the American Geriatrics Society*, 43, 1007–1015.
- Hanson, R. K., & Bussire, M. T. (1998). Predicting relapse: A meta-analysis of sexual offender recidivism studies. *Journal of Consulting and Clinical Psychology*, 66, 348–362.
- Hanson, W. E., Claiborn, C. D., & Kerr, B. (1997). Differential effects of two test interpretation styles in counseling: A field study. *Journal of Counseling Psychology*, 44, 400–405.
- Haynes, S. N., Leisen, M. B., & Blaine, D. (1997). The design of individualized behavioral treatment programs using functional analytic clinical case models. *Psychological Assessment*, 9, 334–348.
- Heaton, R. K., Chelune, G. J., & Lehman, R. A. (1978). Using neuropsychological and personality tests to assess the likelihood of patient employment. *Journal of Nervous and Mental Diseases*, 166, 408–416.
- Herbert, T. B., & Cohen, S. (1993a). Depression and immunity: A meta-analytic review. *Psychological Bulletin*, 113, 472–486.
- Herbert, T. B., & Cohen, S. (1993b). Stress and immunity: A meta-analytic review. *Psychosomatic Medicine*, 55, 364–379.
- Hilsenroth, M. J., Fowler, J. C., & Padawer, J. R. (1998). The Rorschach Schizophrenia Index (SCZI): An examination of reliability, validity, and diagnostic efficiency. *Journal of Personality Assessment*, 70, 514–534.
- Hilsenroth, M. J., Fowler, J. C., Padawer, J. R., & Handler, L. (1997). Narcissism in the Rorschach revisited: Some reflections upon empirical data. *Psychological Assessment*, 9, 113–121.
- Hirschfeld, R. M. A., Klerman, G. L., Andreason, N. C., Clayton, P. J., & Keller, M. B. (1986). Psycho-social predictors of chronicity in depressed patients. *British Journal of Psychiatry*, 148, 648–654.
- Holden, G. (1991). The relationship of self-efficacy appraisals to subsequent health related outcomes: A meta-analysis. *Social Work in Healthcare*, 16, 53–93.
- Hough, L. M. (1992). The “Big Five” personality variables—construct confusion: Description versus prediction. *Human Performance*, 5, 139–155.
- Hunter, J. E. (1986). Cognitive ability, cognitive aptitudes, job knowledge, and job performance. *Journal of Vocational Behavior*, 29, 340–362.
- Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin*, 96, 72–98.
- Kaufman, A. S. (1990). *Assessing adolescent and adult intelligence*. Boston: Allyn & Bacon.
- Kopta, S. M., Howard, K. I., Lowry, J. L., & Beutler, L. E. (1994). Patterns of symptomatic recovery in psychotherapy. *Journal of Consulting and Clinical Psychology*, 62, 1009–1016.
- LaBarbera, J. D., & Cornsweet, C. (1985). Rorschach predictors of therapeutic outcome in a child psychiatric inpatient service. *Journal of Personality Assessment*, 49, 120–124.
- Leavitt, F., & Labott, S. M. (1997). Criterion-related validity of Rorschach analogues of dissociation. *Psychological Assessment*, 9, 244–249.
- Leifer, M., Shapiro, J. P., Martone, M. W., & Kassem, L. (1991). Rorschach assessment of psychological functioning in sexually abused girls. *Journal of Personality Assessment*, 56, 14–28.
- Lezak, M. D. (1995). *Neuropsychological assessment* (3rd ed.). New York: Oxford University Press.
- Loeb, P. A. (1996). *Independent living scales: Manual*. San Antonio, TX: Psychological Corporation.
- LoGiudice, D., Waltrowicz, W., Ames, D., Brown, K., Burrows, C., & Flicker, L. (1997). Health care costs of people referred to an aged care assessment team: The effects of cognitive impairment. *Australian and New Zealand Journal of Public Health*, 21, 311–315.
- Luborsky, L., Chandler, M., Auerbach, A. H., Cohen, J., & Bachrach, H. (1971). Factors influencing the outcome of psychotherapy: A review of the quantitative research. *Psychological Bulletin*, 75, 145–185.
- Luborsky, L., Digner, L., Luborsky, E., McLellan, A. T., Woody, G., & Alexander, L. (1993). Psychological health-sickness (PHS) as a predictor of outcomes in dynamic and other psychotherapies. *Journal of Consulting and Clinical Psychology*, 61, 542–548.
- Magnus, K., Diener, E., Fujita, F., & Pavot, W. (1993). Extraversion and neuroticism as predictors of objective life events: A longitudinal analysis. *Journal of Personality and Social Psychology*, 65, 1046–1053.
- Mattison, R. E., Handford, H. A., Kales, H. C., Goodman, A. L., & McLaughlin, R. E. (1990). Four-year predictive value of the Children's Depression Inventory. *Psychological Assessment*, 2, 169–174.
- McClelland, D. C. (1989). Motivational factors in health and disease. *American Psychologist*, 44, 675–683.
- McKay, J. R. (1991). Assessing aspects of object relations associated with immune function: Development of the Affiliative Trust–Mistrust coding system. *Psychological Assessment*, 3, 641–647.
- Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Kubiszyn, T. W., Moreland, K. L., Eisman, E. J., & Dies, R. R. (1998). *Benefits and costs of psychological assessment in healthcare delivery: Report of the Board of Professional Affairs Psychological Assessment Work Group, Part I*. Washington, DC: American Psychological Association.
- Meyer, G. J., Finn, S. E., Eyde, L., Kay, G. G., Moreland, K. L., Dies, R. R., Eisman, E. J., & Kubiszyn, T. W. (1999). *Psychological testing*

- and psychological assessment: A review of evidence and issues. Manuscript submitted for publication.
- Meyer, G. J., & Handler, L. (1997). The ability of the Rorschach to predict subsequent outcome: A meta-analysis of the Rorschach Prognostic Rating Scale. *Journal of Personality Assessment*, 69, 1-38.
- Moffett, L. A., Steinberg, S. L., & Rohde, P. (1996). Personality assessment of substance-dependent patients in a therapeutic community. *Journal of Substance Abuse Treatment*, 13, 127-134.
- Mohr, D. C., Beutler, L. E., Engle, D., Shoham-Salomon, V., Bergan, J., Kaszniak, A. W., & Yost, E. B. (1990). Identification of patients at risk for nonresponse and negative outcome in psychotherapy. *Journal of Consulting and Clinical Psychology*, 58, 622-628.
- Moras, K. (1997). Toward a core battery for treatment efficacy research on mood disorders. In H. H. Strupp, L. M. Horowitz, & M. J. Lambert (Eds.), *Measuring patient changes* (pp. 301-338). Washington, DC: American Psychological Association.
- Newman, M. L., & Greenway, P. (1997). Therapeutic effects of providing MMPI-2 test feedback to clients at a university counseling service: A collaborative approach. *Psychological Assessment*, 9, 122-131.
- Newnan, O. S., Heaton, R. K., & Lehman, R. A. (1978). Neuropsychological and MMPI correlates of patients' future employment characteristics. *Perceptual Motor Skills*, 46, 635-642.
- Nigg, J. T., Lohr, N. E., Westen, D., Gold, L. J., & Silk, K. R. (1992). Malevolent object representations in borderline personality disorder and major depression. *Journal of Abnormal Psychology*, 101, 61-67.
- Olea, M. M., & Ree, M. J. (1994). Predicting pilot and navigator criteria: Not much more than *g*. *Journal of Applied Psychology*, 79, 845-851.
- O'Leary, A. (1990). Stress, emotion, and human immune function. *Psychological Bulletin*, 108, 363-382.
- Parker, K. (1983). A meta-analysis of the reliability and validity of the Rorschach. *Journal of Personality Assessment*, 47, 227-231.
- Parker, K. P., Hanson, R. K., & Hunsley, J. (1988). MMPI, Rorschach, and WAIS: A meta-analytic comparison of reliability, stability, and validity. *Psychological Bulletin*, 103, 367-373.
- Perry, W., McDougall, A., & Viglione, D. J. (1995). A five-year follow-up on the temporal stability of the Ego Impairment Index. *Journal of Personality Assessment*, 64, 112-118.
- Perry, W., Moore, D., & Braff, D. (1995). Gender differences on thought disturbance measures among schizophrenic patients. *American Journal of Psychiatry*, 152, 1298-1301.
- Perry, W., & Viglione, D. J. (1991). The Ego Impairment Index as a predictor of outcome in melancholic depressed patients treated with tricyclic antidepressants. *Journal of Personality Assessment*, 56, 487-501.
- Piotrowski, C., Belter, R. W., & Keller, J. W. (1998). The impact of "managed care" on the practice of psychological testing: Preliminary findings. *Journal of Personality Assessment*, 70, 441-447.
- Piper, W. E. (1994). Client variables. In A. Fuhrman & G. M. Burlinjame (Eds.), *Handbook of group psychotherapy: An empirical and clinical synthesis* (pp. 83-113). New York: Wiley.
- Pollak, J. M. (1988). The feedback process with parents in child and adolescent psychological assessment. *Psychology in the Schools*, 25, 143-153.
- Prigatano, G. P., Fordyce, D. J., Zeiner, H. K., Roueche, J. R., Pepping, M., & Wood, B. C. (1984). Neuropsychological rehabilitation after closed head injury in young adults. *Journal of Neurology, Neurosurgery, and Psychiatry*, 47, 505-513.
- Project MATCH Research Group. (1997). Matching alcoholism treatments to client heterogeneity. Project MATCH posttreatment drinking outcomes. *Journal of Studies on Alcohol*, 58, 7-29.
- Quirk, M. P., Storsahl, K., Kreilkamp, T., & Erdberg, P. (1995). Personality feedback consultation to families in a managed mental health care practice. *Professional Psychology: Research and Practice*, 26, 27-32.
- Ree, M. J., Earles, J. A., & Teachout, M. (1994). Predicting job performance: Not much more than *g*. *Journal of Applied Psychology*, 79, 518-524.
- Robertson, I. T., & Kinder, A. (1993). Personality and job competencies: The criterion-related validity of some personality variables. *Journal of Occupational and Organizational Psychology*, 66, 225-244.
- Roca, R. P., Klein, L. E., & Vogelsang, G. (1982). Inaccuracy in diagnosing dementia among medical inpatients. *Clinical Research*, 30, 305A.
- Russell, E. W. (1995). The accuracy of automated and clinical detection of brain damage and lateralization in neuropsychology. *Neuropsychology Review*, 5, 1-68.
- Salekin, R. T., Rogers, R., & Sewell, K. W. (1996). A review and meta-analysis of the Psychopathy Checklist and Psychopathy Checklist-Revised: Predictive validity of dangerousness. *Clinical Psychology: Science and Practice*, 3, 203-215.
- Sappington, A. A., & Michaux, M. H. (1975). Prognostic patterns in self-report, relative report, and professional evaluation measures for hospitalized and day-care patients. *Journal of Consulting and Clinical Psychology*, 43, 904-910.
- Saravay, S. M., Steinberg, M. D., Weinschel, B., Pollack, S., & Aloviss, N. (1991). Psychological comorbidity and length of stay in the general hospital. *American Journal of Psychiatry*, 148, 324-329.
- Sbordone, R. J., & Long, C. L. (Eds.). (1996). *Ecological validity of neuropsychological testing*. Delray Beach, FL: St. Lucie Press.
- Schmitt, N., Gooding, R. Z., Noc, R. A., & Kirsch, M. (1984). Metaanalyses of validity studies published between 1964 and 1982 and the investigation of study characteristics. *Personnel Psychology*, 37, 407-422.
- Schulze, D. (1996). Tailor-made and standardized therapy: Complementary tasks in behavior therapy. A contrarian view. *Journal of Behavior Therapy and Experimental Psychiatry*, 27, 119-126.
- Schwartz, S., & Wiedel, T. C. (1981). Incremental validity of the MMPI in neurological decision-making. *Journal of Personality Assessment*, 45, 424-426.
- Shapiro, P. A., Williams, D. L., Foray, A. T., Gelman, I. S., Wukich, N., & Sciacca, R. (1995). Psychosocial evaluation and prediction of compliance problems and morbidity after heart transplantation. *Transplantation*, 60, 1462-1466.
- Snyder, D. K., Mangrum, L. F., & Wills, R. M. (1993). Predicting couples' response to marital therapy: A comparison of short- and long-term predictors. *Journal of Consulting and Clinical Psychology*, 61, 61-69.
- Spengler, W. D. (1992). Validity of questionnaire and TAT measures of need for achievement: Two meta-analyses. *Psychological Bulletin*, 112, 140-154.
- Stricker, G., & Healey, B. J. (1990). Projective assessment of object relations: A review of the empirical literature. *Psychological Assessment*, 2, 219-230.
- Teitelbaum, L. M., & Carey, K. B. (1996). Alcohol assessment in psychiatric patients. *Clinical Psychology: Science and Practice*, 3, 323-338.
- Tessler, R., Mechanic, D., & Dimond, M. (1976). The effect of psychological distress on physician utilization: A prospective study. *Journal of Health and Social Behavior*, 17, 353-364.
- Tett, R. P., Jackson, D. N., & Rothstein, M. (1991). Personality measures as predictors of job performance: A meta-analytic review. *Personnel Psychology*, 44, 703-742.
- Thomas, C. B. (1988). Cancer and the youthful mind: A forty-year perspective. *Advances*, 5, 42-58.
- Tuber, S. B. (1983). Children's Rorschach scores as predictors of later adjustment. *Journal of Consulting and Clinical Psychology*, 51, 379-385.
- Tuber, S. (1992). Empirical and clinical assessments of children's object relations and object representations. *Journal of Personality Assessment*, 58, 179-197.
- Urist, J. (1977). The Rorschach test and the assessment of object relations. *Journal of Personality Assessment*, 41, 3-9.

Urist, J., & Shill, M. (1982). Validity of the Rorschach mutuality of autonomy scale: A replication using excerpted responses. *Journal of Personality Assessment*, *46*, 450-454.

Verive, J. M., & McDaniel, M. A. (1996). Short-term memory tests in personnel selection: Low adverse impact and high validity. *Intelligence*, *23*, 15-32.

Weiner, I. B. (1994). Rorschach assessment. In M. E. Maruish (Ed.), *The use of psychological testing for treatment planning and outcome assessment* (pp. 249-278). Hillsdale, NJ: Erlbaum.

Weiner, I. B. (1996). Some observations on the validity of the Rorschach Inkblot Method. *Psychological Assessment*, *8*, 206-213.

Weiner, I. B., & Exner, J. E., Jr. (1991). Rorschach changes in long-term and short-term psychotherapy. *Journal of Personality Assessment*, *56*, 453-465.

Weissman, M. M., Prusoff, B. A., & Klerman, G. L. (1978). Personality and the prediction of long-term outcome of depression. *American Journal of Psychiatry*, *135*, 797-800.

Wells, K., Stewart, A. L., Hays, R. D., Burnam, A., Rogers, W. H., Daniels, M., Berry, S. D., Greenfield, S., & Ware, J. E. (1989). The functioning and well-being of depressed patients: Results from the Medical Outcomes Study. *JAMA: Journal of the American Medical Association*, *262*, 914-919.

Wetzler, S., & Marlowe, D. B. (1993). The diagnosis and assessment of depression, mania, and psychosis by self-report. *Journal of Personality Assessment*, *60*, 1-31.

Wilson, B. A., Cockburn, J., Baddeley, A. D., & Hiorns, R. (1989). The development and validation of a test battery for detecting and monitoring everyday memory problems. *Journal of Clinical and Experimental Neuropsychology*, *11*, 855-870.

Wilson, B. A., Cockburn, J., & Halligan, P. (1987). *The Behavioral Inattention Test*. Flemspton, Bury St. Edmunds, Suffolk, England: Thames Valley Test Company.

Zalewski, C. E., & Gottesman, I. I. (1991). (Hu)man versus mean revisited: MMPI group data and psychiatric diagnosis. *Journal of Abnormal Psychology*, *100*, 562-568.

Received April 5, 1999

Revision received November 15, 1999

Accepted December 2, 1999 ■



**AMERICAN PSYCHOLOGICAL ASSOCIATION  
SUBSCRIPTION CLAIMS INFORMATION**

Today's Date: \_\_\_\_\_

We provide this form to assist members, institutions, and nonmember individuals with any subscription problems. With the appropriate information we can begin a resolution. If you use the services of an agent, please do NOT duplicate claims through them and directly to us. **PLEASE PRINT CLEARLY AND IN INK IF POSSIBLE.**

PRINT FULL NAME OR KEY NAME OF INSTITUTION	MEMBER OR CUSTOMER NUMBER (MAY BE FOUND ON ANY PAST ISSUE LABEL)
ADDRESS	DATE YOUR ORDER WAS MAILED (OR PHONED)
CITY STATE/COUNTRY ZIP	<input type="checkbox"/> PREPAID <input type="checkbox"/> CHECK <input type="checkbox"/> CHARGE CHECK/CARD CLEARED DATE: _____
YOUR NAME AND PHONE NUMBER	(If possible, send a copy, front and back, of your cancelled check to help us in our research of your claim.)
TITLE	ISSUES: <input type="checkbox"/> MISSING <input type="checkbox"/> DAMAGED
	VOLUME OR YEAR
	NUMBER OR MONTH

*Thank you. Once a claim is received and resolved, delivery of replacement issues routinely takes 4-6 weeks.*

<b>(TO BE FILLED OUT BY APA STAFF)</b>	
DATE RECEIVED: _____	DATE OF ACTION: _____
ACTION TAKEN: _____	INV. NO. & DATE: _____
STAFF NAME: _____	LABEL NO. & DATE: _____

Send this form to APA Subscription Claims, 750 First Street, NE, Washington, DC 20002-4242

**PLEASE DO NOT REMOVE. A PHOTOCOPY MAY BE USED.**