Visions of Clinical Assessment: Then, Now, and a Brief History of the Future

Gary Groth-Marnat
Curtin University

Perspectives on the discipline of psychological assessment are provided through reflections on a classic article by Hunt (1946), review of current issues, and a projection of changes likely to occur over the next fifty years. Topics covered include the place of idiosyncratic qualitative responses, test design, clinical judgment, managed care, financial efficacy, over emphasis on pathology, treatment planning, status of projectives, ecological validity, relationship between theory and tests, innovations in scale construction, and computer-assisted assessment. Future predictions relate to innovations in computer-assisted assessment, practitioner roles in an increasingly automated environment, the future of traditional tests, and human-rights issues. © 2000 John Wiley & Sons, Inc. J Clin Psychol 56: 349–365, 2000.

Psychological assessment always has been one of the major core-defining features of clinical psychology (Garfield & Kurtz, 1973; Piotrowski & Zalewski, 1993, Weiner, 1983). The arrival of the year 2000 invites us as a profession to reflect upon the past, present, and future of clinical assessment. Such an evaluation optimally will help to crystallize its strengths and weaknesses and set an agenda for the next fifty years and beyond.

Research on the extent and patterns of test usage indicate that there have been a number of noteworthy changes over time. During the birth and expansion of clinical psychology directly after World War II, psychologists spent nearly half their time conducting assessments (Lubin, Larsen, & Matarazzo, 1984; Lubin, Larsen, Matarazzo, & Seever, 1985, 1986; Sundberg, 1961). Much of the focus was on using the Rorschach along with the newly developed MMPI and Wechsler Bellevue scales. Over the past 50 years, the overall time spent in assessment gradually has decreased to 28% in 1969, 22% in 1982 (Lubin et al., 1984, 1985, 1986), and 20% in 1995 (Watkins, Campbell, Nieberding, & Hallmark, 1995). Despite this decrease, Watkins et al. (1995) noted that 91% of

Correspondence concerning this article should be addressed to Gary Groth-Marnat, School of Psychology, Curtin University, GPO Box U1987, Perth, WA 6001 Australia; email: garygm@psychology.curtin.edu.au

JOURNAL OF CLINICAL PSYCHOLOGY, Vol. 56(3), 349–365 (2000)

all practicing psychologists engaged in at least some form of assessment, and Kinder (1994) noted that 64% of all nonacademic job advertisements listed assessment as an important prerequisite. Despite these data suggesting continued widespread use, the current managed-care environment is resulting in apparent reductions in assessment (Piotrowski, 1999) primarily due to restrictions in reimbursement (Stout & Cook, 1999).

Over 50 years ago, Hunt (1946), in his "The future of diagnostic testing in clinical psychology" (Journal of Clinical Psychology, Vol. 2, No. 4), reflected on both the patterns and trends that he observed in assessment and attempted to assert values and strategies relevant for its future. Specifically, Hunt emphasized that clinical psychologists should pay more attention to qualitative behavior during the testing situation, rework tests to yield a maximum amount of rich qualitative responses, emphasize the importance of clinical judgment, and study clinical judgment in much the same way as a formal psychological test. It is crucial to review each one of these points in terms of how they have had an impact on the practice and training of clinical psychology over the past 50 plus years. Such a review will help to extract which contemporary issues the field is confronting, which will then enable a projection ahead into the next 50 years.

The Importance of Qualitative Behavioral Observations

Hunt (1946) expressed concern that numerical units (IQ, scaled scores, percentiles, etc.) seemed to be given too much emphasis compared to rich qualitative behaviors. He provided examples of different types of responses to the Wechsler intelligence scales, including a tangential response by a schizophrenic and an apologetic response by a seeming perfectionist. These responses might result in the same overall score as other clients, but the reason *why* they arrived at the correct (or incorrect) response might be quite different. However, only attending to the score might obscure clinically useful information. Most clinicians similarly can recall anecdotal situations relating to the importance of attending to qualitative observations. For example, one client, when asked what to do with a self-addressed stamped enveloped he had found, stated that he would soak the stamp off. The client never paid his bill (although this outcome might have been as easily predicted by noting the quite high score on his MMPI Pd scale).

In stressing the above dilemma (and Hunt's value in terms of what he feels would be most clinically relevant), Hunt has outlined nicely a controversy that has emerged in various forms and been researched extensively over the past 50 years. Specific variations of the controversy can be seen in the relative accuracy of clinical versus actuarial prediction, research on clinical judgment, the role of tester as technician (psychometrist) versus problem-solving clinician, nomothetic versus idiographic emphasis, and the pathognomonic sign versus the quantitative fixed-battery approach in neuropsychology. Hunt clearly has placed himself in the qualitative end of the continuum. This is not to say that he also would not have used and believed in quantitative scores. However, he does imply that the stagnation he felt as a part of assessment in 1946 partially could be attributed to not giving sufficient consideration to nuances of behavior by the client. Accordingly, he encourages that both test design and clinical training should be organized to be more sensitive to such qualitative observations.

The above controversy is far from solved and each professional psychologist varies somewhere on the qualitative/quantitative continuum. There is also general agreement as to the importance of both of these avenues of understanding the client. Note, for example, that nearly all Wechsler interpretive approaches recommend attending to a client's idiosyncratic responses to selected items. A qualitative approach to the MMPI-2/MMPI-A also can be found when exploring the meanings behind a client's responses to critical

items. In addition, a large proportion of current clinical neuropsychologists identify with strategies that emphasize understanding the reasons why a client arrived at a particular score (Sweet, Moberg, & Westergaard, 1996). Alexander Luria, Edith Kaplan, and Kevin Walsh each have emphasized and developed various strategies to understand better the processes clients go through when developing responses (Kaplan, Fein, Morris, & Delis, 1991; Groth-Marnat, Gallagher, Hale, & Kaplan, 2000). This has influenced sufficiently the development of the WAIS-III so that procedures can be used on Digit-Symbol Coding to determine whether a moderate score was due mainly to speed versus memory. In contrast, others have argued that a more quantitative or fixed-battery approach has the advantage of providing a constant background pattern of results, and is easier to research, teach, standardize, and validate (Russell, 1998).

There also have been considerable challenges to qualitative, flexible, clinician-oriented approaches. It has been argued that clinicians are likely to *overemphasize* single, dramatic, "pathognomonic" signs (Wedding & Faust, 1989) and *de-emphasize* bland information, such as quantitative scores or employment history, even when the more bland, quantitative information produces more accurate judgments. The result is likely to be judgment errors when clinicians are exposed to idiosyncratic sources of information. This is consistent with Sawyer's (1966) classic review of the literature, which noted that judgment accuracy frequently *decreased* when interview information was combined with test data compared with when test data is used on its own. In addition, clinicians seem likely to seek, use, and recall information that is likely to confirm but not to refute their initial impressions (confirmatory hypothesis testing; Faust, 1991; Garb, 1998, pp. 59–83; Strohmer & Newman, 1983; Wedding & Faust, 1989). If the clinicians were given different types of initial information or for other reasons developed different initial hypotheses, they might have arrived at quite different conclusions.

It seems that the nature of progress in science is that people initially take sides in a position, divide into warring camps ("hard-nosed skeptical academics versus soft-hearted gullible practitioners"), and eventually integrate the opposing positions into a more inclusive perspective. Note, for example, how hypnosis theorists at one time were divided between hypnosis being an altered state versus a socially constructed process. Progressively, more professionals now see that these can be complementary rather than competing ways of understanding hypnotic processes. Similarly, behaviorism's turn inward to cognitions, and especially schemas, has brought it to philosophically and strategically more similar psychoanalytic approaches. It is likely that advances in the field of assessment also will benefit from further complementarity and integration. For example, Hammainen (1994) has developed a computer-assisted expert system for neuropsychological assessment that integrates the use of quantitative and qualitative procedures. In addition, the current computer-scoring system for the California Verbal Learning Test provides both quantitative scores as well as scoring for more qualitative responses (primacy versus recency effects, intrusions). Such qualitative scoring may help determine the possibility of a client malingering or experiencing specific types of brain impairment (Fridlund & Delis, 1987).

Tests Should be Designed to Yield Both Rich Observable Behavior and Convenient Numerical Scores

During the 1940s and 1950s, there was a positive consensus on the part of clinical psychologists of the value of projective tests along with faith in a clinician's ability to use effectively this information. Hunt reflects this in his belief in the crucial value of idiosyncratic responses to the Wechsler Bellevue and other forms of assessment: "Let us

select items not only to allow a numerical measure, but also to provide the subject with an opportunity for revelatory clinical behavior even though it goes beyond the present potentialities of objectification." (p. 313). This desire to maximize qualitative information also was reflected in the field in general. At the same time Hunt's article was published, the second volume of Rapaport, Gill, and Schafer's (1946) *Diagnostic Psychological Testing* became available and rapidly became the defining resource in assessment. A survey of its contents indicates an emphasis on projective tests (Rorschach, TAT), psychoanalytic theory, and a focus on idiosyncratic features of the client.

Hunt's suggestion and wish for tests to emphasize items that would pull for rich observable behavior mostly has gone unrealized for several reasons. First, there have been consistent, and sometimes passionate, attacks on the psychometric properties of projective techniques. For example, Garb (1998, pp. 57–59, 237) summarized the incremental validity of projectives by stating that their inclusion with other test and demographic information typically *decreased* the accuracy of clinical judgments when compared to judgments made with the same information but without projectives. In addition, the one-time hegemony of psychoanalysis has been supplanted by briefer, focused, more scientifically verifiable methods of assessment and intervention. While rich behavioral observations certainly can occur outside the context of psychoanalysis, it was/is psychoanalysis that typically has championed and emphasized the importance of idiosyncratic qualitative information. Finally, the streamlined, cost-conscious managed-care environment is likely to be skeptical of the cost benefits of the more labor-intensive techniques typical of projectives. Current thinking is that indeed idiosyncratic qualitative responses can be important, but this importance is primarily in generating hypotheses that need to be verified by a number of additional sources and strategies. However, inferences derived from idiosyncratic behavioral information need to be made cautiously and strategies employed (i.e., considering possible refutations of hypotheses) to guard against clinician bias.

One means of focusing on a client's uniqueness has been behavioral assessment. Unlike traditional assessment, which would use a client's idiosyncratic behavior to infer internal traits or abilities, behavioral assessment has focused on situational determinants, especially antecedents and consequences of behavior. Both Hunt (1946) and proponents of behavioral assessment might have agreed that clinical assessment was in a state of stagnation, but for different reasons. Hunt emphasized assessment's suboptimal focus on idiosyncratic behavior, but he was still interested in how these behaviors could be used to infer internal client characteristics. In contrast, behavioral assessment has criticized traditional assessment as being too focused on unobservable behaviors that are not connected sufficiently to interventions. Behavioral assessment historically has felt that traditional psychology has stagnated because interventions were overly interested in verbal techniques and were not sufficiently powerful. Over the past decade, however, traditional and behavioral assessments have become aligned more closely. Specifically, behavioral assessment has turned more inward towards evaluating cognitions and more outward in its interest in evaluating the psychometric properties of its techniques, as well as recognizing the relevance of DSM diagnosis (Groth-Marnat, 1999). Conversely, traditional assessment has been replaced to some extent by an emphasis on an objective, problem-focused, self-report approach (e.g., anger, hostility, anxiety).

Emphasize and Study Clinical Judgment

Hunt (1946) both expressed faith in the ability of clinicians to make accurate judgments and emphasized that these processes should be thoroughly studied. In so doing, he antici-

pated the great and longstanding debate between clinical as opposed to actuarial (statistical) accuracy. Just eight years after Hunt's article, Paul Meehl (1954) published his now classic *Clinical versus Statistical Prediction: A Theoretical Analysis and Review of the Evidence*, which surveyed the current research and concluded that actuarial judgment was generally more accurate than judgments made by clinicians. Over the next 50 years, this stimulated lively debate, as well as literally thousands of relevant research publications. This research generally has supported Meehl's initial conclusions.

Despite general empirical support for the greater accuracy of actuarial techniques, there are a number of practical and theoretical issues that underscore Hunt's emphasis on the role of the clinician. In some contexts and for some types of data, a flexible clinical interview is the only means of obtaining the required information. Unusual events or qualitative behavioral observations often can result in modifying a clinician's conclusions. In addition, an assessment is typically client and context specific in that conclusions need to be directed towards solving a specific type of problem. Actuarial formulas typically do not take into account either the client's context or the complexities of how the information will be used. Similarly, actuarial approaches conceptualize the world as stable. In contrast, practitioners must deal with a fluid, imperfect, changing world that is filled with changing perceptions and in which it is difficult to predict chance events (Groth-Marnat, 1999). This means that the clinician might be advised to use actuarial methods whenever possible, but he or she still needs to use this information in a flexible manner to deal with the unique context of a client's presenting problem.

The most complete and thorough review of clinical judgment can be found in Garb's (1998) *Studying the Clinician: Judgement Research and Psychological Assessment*. One of the major conclusions is that the accuracy of clinical judgment often is questionable. Errors can occur from a wide number of areas, including not taking into consideration base-rate frequencies, primacy effects, confirmatory bias, hindsight bias, attribution bias, clinician personality type (i.e., authoritarian personality), and dissimilarity between clinician and client. One finding has been that, for some areas, lay interviewers were as accurate as trained clinicians. In contrast, trained clinicians were better than lay interviewers for more complex tasks, such as making clinical diagnoses; rating mental status; short-term prediction of violence, competency evaluations, and interpreting psychological tests, psychotherapy data, and biographical information. Interestingly, confidence in the accuracy of judgements generally was not related to accuracy unless the clinician was extremely expert in the specific area that they were judging. Fortunately, the accuracy of clinical judgment can be increased through a variety of strategies, including the use of comprehensive interviews, taking into consideration arguments that might not support clinician inferences, strong reliance on carefully written notes (rather than memory), attention to base rates, seeking feedback on accuracy, increased knowledge related to the domain of behaviors being judged, and reliance on specific DSM-IV criteria when making diagnoses.

The very brief outline of clinical judgment mentioned above indicates that, in some ways, Hunt's (1946) wish for research on clinical judgment partially has been fulfilled. However, Hunt also suggested that research should be focused on the individual clinician to determine who was accurate in some areas or client types and who was not. The result then would either be enhanced training to improve accuracy or a restriction of practice only to areas in which the clinician was accurate. For the most part, the use of judgment research for guiding practice and training has not been implemented to the extent that it should or could be. Garb (1998), for example, has lamented that ". . . one of the main findings of judgment research is that clinicians do not attend closely to empirical findings . . ." (p. 243). He further notes that effective and ineffective aspects of training programs

should be identified, licensing should be made more relevant to judgment tasks, and instructional aids should be developed and used. The above "unfinished" but applied aspects of judgment research thus represent a future agenda for clinical psychology.

Additional Current Issues in the Field

Hunt's article extracted and set the stage for a number of crucial areas of research and controversy for the next 50 years. However, he focused primarily on the issue of idiosyncratic behavioral observations and clinical judgment. While these concerns are still important, there are a number of additional issues currently confronting psychological assessment. Some of these have come from sources external to the field, including general skepticism (and little reimbursement) from managed health care, as well as a feeling among many clients and psychologists that clinical assessment focuses too much on negative qualities. There also is criticism from within the field pointing out that assessment is not cost effective. It also overemphasizes pathology, does not pay sufficient attention to treatment planning, does not relate sufficiently to everyday types of client behaviors (ecological validity), has not sufficiently incorporated new theories, and has not addressed the continued criticism and ambivalence towards projective techniques. In addition, the full potential of computer technology has not been realized in that the focus has been primarily on their clerical and interpretive capacities rather than novel presentation of stimuli, integration of psychophysiological measures, and artificial intelligence.

Managed Care

The greatest current threat to assessment comes from managed care. Many managed-care companies perceive clinical assessment, especially the full battery, as neither crucial to treatment planning or other forms of decision making nor cost effective (Eisman et al., 1998; Griffith, 1997). Accordingly, Stout and Cook (1999) found that 30% of managed-care organizations did not provide any reimbursement for formal assessment and the rest provided varying levels of reimbursement. It is rare for a full psychological battery to be reimbursed, and time spent in report writing, client feedback, and discussions with referral sources also typically are not reimbursed (Eisman et al., 1998; Griffith, 1997). Anecdotal reports also suggest that many policy makers consider psychological testing to have been a sometimes overused "cash cow" for providers in the past; this position is one reason for justifying reduced reimbursement.

The result of the above trend is that there has been a significant reduction in the amount of time professional psychologists spend conducting assessment (Piotrowski, 1999). When assessment is required, it is likely to rely on a combination of clinical interview and brief symptom-oriented measures that target such areas as depression, anxiety, and anger. In contrast, time-consuming instruments such as the Wechsler intelligence scales and Rorschach are being used less frequently. However, when time or cost is not an issue, testing practices seem to be occurring in much the way that they have in the past. This typically includes forensic settings, neuropsychological contexts, personnel selection, and educational settings where government mandates require comprehensive assessment (Piotrowski, 1999).

Several recent studies on the impact of managed health care on graduate training in psychological assessment have appeared. Piotrowski and Belter (1999) reported data from 84 Directors of APA-approved internship training programs on changes in assess-

ment training. Interestingly, less than half of the programs indicated that emphasis in psychological assessment significantly has been affected by managed care, suggesting that most sites are somewhat insulated from current marketplace pressures. However, there was a slight decline in assessment in general, particularly with projective techniques. Compared to prior studies in this area, the majority of programs have continued to stress objective measures, IQ tests, interview techniques, and behavioral approaches to testing. In a related work, these same authors conducted a study of training trends in testing at master's degree programs and found a marked decrease in emphasis on projective techniques in recent years. However, traditional methods of assessment were still being taught and competency in overall assessment skills continues to be considered a central component of graduate training (Belter & Piotrowski, in press).

The above nexus presents a number of challenges to the profession. New niches might be developed in such areas as prevention, chronic pain, somatization, geriatrics, and behavioral dentistry (Stout & Cook, 1999). In addition, psychologists need to place greater efforts into demonstrating the value of clinical assessment (Meyer et al., 1998). Such value might be in the form of accurate and rapid diagnosis (especially when complex differential diagnosis is an issue), designing optimal treatment plans, monitoring client progress, and evaluating treatment outcome. It would seem that, given managed care's emphasis on evidence-based procedures, psychological assessment should be tied closely to treatment planning and outcome. Hopefully, future health-care standards will demand that the assessment–treatment link will be integrated more tightly and progressively more dependent on empirical evidence.

Financial Efficacy

At the core of managed health care's concerns with assessment is that it is not financially efficient. Unfortunately, little research has been conducted to demonstrate whether, or under which circumstances, assessment is financially efficacious. In contrast, research in industrial/organizational settings have been concerned far more with the financial implications of testing procedures. In order to optimize the financial efficacy of clinical assessment, the following rational strategies (see Groth-Marnat, 1999) might be utilized: assess domains most relevant to treatment planning/outcome; use formal assessment when indicated to reduce risk management; target conditions most likely to realize cost savings (i.e., panic, somatization); increase automated procedures; use time-efficient instruments; develop closer links between assessment, feedback, and therapy; and integrate treatment planning, progress monitoring, and outcome evaluating. In addition, a clear agenda for professional psychology should be to demonstrate more clearly the financial efficacy of clinical-assessment methods.

Overemphasis on Pathology

It is nearly a cliche that clinical psychology has adopted the medical model's emphasis on identifying problem areas without paying sufficient attention to a client's positive qualities. Interpretations of results from the MMPI-2/MMPI-A and MCMI-III often can seem like wading through a cesspool of psychopathology. The items themselves represent this "searching-for-pathology" model, and clients frequently comment on this feature. The result is that clients, who might otherwise be advocates of testing, are somewhat wary of the process. While there have been some tests designed to assess self-esteem, wellness, or self-actualization, these are certainly not in widespread clinical use. It could be argued

that none of the 38 tests listed in the Watkins et al. (1995) survey of test-use patterns specifically focus on client strengths. However, there are a number of infrequently used tests that are either pathology neutral or measure bipolar dimensions of strengths and weaknesses (Vineland Social Maturity Scale, Strong Interest Inventory, 16PF, California Psychological Inventory, Edwards Personal Preference Inventory ranked 16th, 18th, 24th, 30th, and 32nd, respectively). Neuropsychological assessment similarly uses assessment tools that are directed almost exclusively towards identifying and elaborating on dysfunction (Groth-Marnat et al., 2000). What ideally is needed is a comprehensive ("multiphasic"), psychometrically sound instrument that assesses prosocial behaviors, coping mechanisms, healthy behaviors, optimal cognitions, and self-efficacy, along with a solid rationale for using it. This then could be used to balance the assessment of problem areas with client strengths. Part of the rationale would be that it would be likely to enhance the clinician–client relationship. The rationale for measuring client strengths also might involve demonstrating that it has treatment and cost utility in that it would increase the rate, as well as optimize the outcome, of treatment. Thus, when a test survey is conducted ten or twenty years from now, it would be hoped that such a test would be in the top-ten most-frequently-used assessment tools.

Treatment Planning

Over the past 50 years, most treatment planning relied on a combination of clinical judgement guided by clinical "lore." Progressively, such efforts have been guided by more empirically validated approaches. Particularly promising areas over the past decade include research on treatment planning based on client characteristics (Beutler, Clarkin, & Bongar, in press; Harkness & Lilienfeld, 1997; Wickramasekera, 1995), empirically validated treatments based on diagnosis (Compass, Haaga, Keefe, Leitenberg, & Williams, 1998; Task Force, 1995), educational interventions tailored around cognitive abilities (Naglieri, 1999), client–therapist matching (i.e., Jacobson, Follette, & Pagel, 1986), and stages of change (Prochaska, DiClemente, & Norcross, 1992). Instruments designed specifically for treatment planning include the *Butcher Psychological Treatment Planning Inventory* (Butcher, 1998) and *Systematic Treatment Selection* (Beutler & Williams, 1998; Fisher, Beutler, & Williams, 1999). The above clearly represents efforts to develop theoretically sound, empirically validated means of treatment planning. However, major impediments to such efforts include acceptance by practitioners and managed-care providers, incorporation into training programs, and the degree of practitioner flexibility in actually changing clinician behaviors to reflect treatment guidelines.

Integration of Assessment and Intervention

Fifty years ago, it was traditional for psychologists to conduct assessments and not share the results with clients. In the past two decades, ethical guidelines (American Psychological Association, 1992) and client-advocacy legislation has made client feedback a standard procedure. Such feedback is expected to be clear, direct, accurate, and in terms understandable to the client (Pope, 1992). Effective feedback also has been found to have direct therapeutic benefits for the client as well (Finn & Martin, 1997; Gass & Brown, 1992).

The Status of Projective Techniques

Despite consistent controversy, projective techniques have shown remarkable resiliency over time. Their attraction seems to be that they provide clinicians with rich, qualitative

data, are perceived as being able to obtain information "beneath" a client's conscious defenses (e.g., covert and deep information), assess global aspects of personality, and often are perceived by clients as intrinsically interesting and nonthreatening. However, there is some indication that recently the use of projectives has been declining, and this is likely to be due to the impact of managed care (Piotrowski, 1999). Managed-care providers typically perceive that projectives are time consuming and provide information that only is marginally useful in decision making, client monitoring, and treatment planning.

Probably the major projective test that is most likely to decline is the Thematic Apperception Test and, to a lesser extent, projective drawings. In contrast, there is still continued interest and research in the Rorschach primarily due to the efforts of Exner and Weiner (Exner, 1993). Even though older and more recent criticisms of the Rorschach linger, meta analyses, primarily using Exner's Comprehensive System, mostly have supported its reliability and validity (Parker, 1983; Parker, Hanson, & Hunsley, 1988). At least part of the attraction of the Rorschach is that it fills a unique niche in offering an understanding of the underlying perceptual processes of schizophrenia and other forms of psychopathology involving disordered thinking and perceptual processes. Until alternative instruments are developed to understand these aspects of functioning, it is likely that the Rorschach will continue to be attractive despite the fact that it is labor intensive and has a number of psychometric shortcomings. It might be speculated that, with the current interest in cognitive processes such as schemas and other "nonconscious" forms of information processing, new assessment strategies might be developed that will compete with the Rorschach. Such techniques would draw heavily on information-processing technologies and computer modeling rather than psychoanalytic theories.

Ecological Validity

Test validity traditionally has not focused on a person's everyday abilities/characteristics, but rather on such areas as correlations with previously developed measures, factor analyses, and level of diagnostic agreement. Practitioners then must make inferences related to work, recreation, relationships, or aspects of educational performance. Often these inferences are not based sufficiently on empirical research. This is consistent with a survey of the usefulness of IQ tests, which found that they were perceived by practitioners as being least appropriate for real-life problem-solving situations (Harrison, Kaufman, Hickman, & Kaufman, 1988). This issue recently has been receiving increased attention, with various reviews urging more extensive work in establishing the everyday (ecological) validity of tests (see Acker, 1990; Groth-Marnat & Teal, in press; Sbordone & Long, 1996). In addition, some tests, such as the Rivermead Behavioral Memory Test (Wilson, Cockburn, & Baddeley, 1985), have been designed to use more everyday test stimuli in an attempt to increase their face, as well as their ecological, validity. A clear agenda for clinical assessment would be to continue making further efforts to enhance the ecological validity of instruments.

Relationship Between Theory and Tests

The relationship between tests and the theories that underlie them often has been tenuous. This particularly is true in the assessment of intelligence, where their main practical purpose historically has been to predict global educational and vocational performance. New developments in intellectual assessment have attempted to tighten this theory–test

relationship (see Harrison, Flanagan, & Genshaft, 1997). For example, Luria's PASS (Planning–Attention–Simultaneous–Successive) model has guided the development and interpretation of the Cognitive Assessment System (Naglieri, 1999), Kaufman Assessment Battery for Children (Kaufman & Kaufman, 1983), Woodcock–Johnson Tests of Cognitive Ability (Woodcock & Johnson, 1989), and the Luria–Nebraska Neuropsychology Battery (Golden, Purisch, & Hammeke, 1985). Another avenue of development, which corresponds to the previous issue of ecological validity, is Sternberg and his colleague's (Sternberg, Wagner, Williams, & Horvath, 1995) interest in developing tests to measure practical, commonsense problem solving, such as the Tacit Knowledge Inventory for Managers (Wagner & Sternberg, 1991). A final example that has inspired many educators and eventually may result in formal, empirically based assessment procedures is Gardner's concept of multiple intelligence (Chen, Isberg, & Krechevsky, 1998; Gardner, 1993).

A larger issue is the underlying model for predicting behavior. For example, Bayesian concepts, such as effect size, positive/negative predictive power, and base rates, have begun to occur more frequently in clinical assessment. A more radical departure from traditional models is the inclusion of chaos theory (Heiby, 1995), which emphasizes the prediction of idiographic, nonlinear, dynamic behavior. For example, assessment inspired by chaos theory might include measures of an individual over a period time (time-series assessment) in which vectors (the often simultaneous impact of several forces) might be combined with understanding a bifurcation (point at which there is some crucial change in behavior). Chaos theory uses mathematical models and emphasizes the interconnectedness of events. Since much of human behavior is considered unstable and difficult to predict, models based on chaos theory might become increasingly attractive for clinical assessment.

Innovations in Scale Construction

Scale development continually has worked to refine more sophisticated methods of item selection, scale construction, and item presentation. The traditional approaches have been through selecting items based on rational (content), empirical, or factor-analytic approaches. Several decades ago, it was possible to find most tests exclusively using one of these approaches. More recent trends indicate the use of a combination of various strategies. For example, the original MMPI typically has been considered the clearest and strongest example of empirical criterion approaches. However, the revised MMPI-2 has used content strategies to develop new scales and to refine the interpretation of the traditional clinical scales. In addition, the MCMI used a combination of rational theory-based, as well as empirical, procedures that have incorporated Bayesian concepts (i.e., use of base rates and positive/negative predictive power). Currently, item-response theory has gained considerable support in that it provides a rigorous analysis of items along with the possibility of adapting the sequence and types of items to the ability level of each client. In other words, by fully understanding the characteristics of each item and having an extensive item pool, items can be given to clients to explore more fully their ability, level of achievement, or personality characteristics. Items that are not likely to contribute much to such an understanding can be omitted. Traditional paper–pencil tests have difficulty utilizing sophisticated item presentation tailored to the level of the person. However, adaptive administrations based on item-response theory can be developed easily and optimally for presentation by computers such that testing efficiency could be increased by 50% or more (Weiss, 1985).

Computer-Assisted Assessment

Computers (and technologies in general) offer wondrous solutions, but, at the same time, they open up new dilemmas for the profession and society. In terms of solutions, many of the previously discussed issues can be addressed. Computers offer strategies of enhancing clinical judgment, improving time (and cost) efficiency, improving ecological validity, optimizing models for treatment planning, and incorporating new assessment theories. Despite this potential, computer technology to date has focused primarily on clerical efficiency, data storage, and generating interpretive hypotheses (McMinn, Buchanan, Ellens, & Ryan, 1999). Most of the commonly used personality measures offer each of these services. Probably the most sophisticated integration of services for traditional tests is the Psychological Corporation's Optaio Network (Psychological Corporation, 1998), which allows users to do on-line assessments, develop summary reports, create treatment plans, and track client outcomes.

Computer-assisted approaches in clinical neuropsychology include full batteries tailored towards detecting the impact of substance abuse, performance of airline pilots, and exposure to neurotoxic substances (Kane & Kay, 1992). These programs primarily are used in large organizational contexts. Some of the tests simply have used computerized standard tests, such as finger tapping or the Stroop technique. Other strategies have capitalized on the computer's flexible, ecologically congruent potential, such as the stimulus presentations of the Memory Assessment Clinic's battery. For example, one of its subtests displays different rooms in a 12-room house, instructs the client to place 20 common objects in different locations of the client's own choosing, and then requires the client to recall the locations of these objects (Larrabee & Crook, 1988). Somewhat similar, Davidson, Stevens, Goddard, Bilkey, and Bishara (1987) describe a measure of attention that uses an animated tracking task, which requires clients to "drive" a car down a road that is scrolled on a screen in front of them. Although such approaches show considerable promise, they neither have begun to replace the more traditional neuropsychological batteries/tests nor do they have near the amount of validation research available.

Despite the above computer-assisted developments, there are still a large number of mostly untapped computer uses. One way of conceptualizing the actual and potential use of computers is to consider the following continuum of complexity beginning with simple clerical efficiency (scoring) and extending to more novel presentation of stimuli and interpretation of responses:

1. *Clerical efficiency*—Scoring and data storage.
2. *Interpretation*—By using either expert derived or actuarial strategies, a much larger base for generating hypotheses can be developed.
3. *Innovative presentation of traditional test items*—Adaptive selection of test items based on item-response theory could increase the efficiency of assessment; based on a client's previous responses, irrelevant items could be skipped or other areas explored in more depth.
4. *Networked norms*—Internet storage and reference to norms might be developed such that norms could be both tailored to an individual client and each new client assessment could be stored in an ever increasing database.
5. *Presentation of novel stimuli*—Unlimited potential for how test "items" might be presented; this might include virtual-reality simulations, which are far more complex, rich, and life-like (hence with high face and ecological validity) than traditional paper–pencil tests; test could be interactive, actual interpersonal or

task-related simulations might be presented; test could be voice activated; complex interpersonal responses might be noted and analyzed.

6. *Time-series measures*—Rather than one or two assessment sessions, computerized recording devices can be attached to the client to measure ongoing sequences of behaviors over a period of days or weeks; such information might be transferred to central data storage and integrated with additional assessment information (test results, demographics, medical records, etc.).
7. *Psychophysiological monitors*—Psychophysiological responses could be integrated with other forms of responses.
8. *Artificial intelligence*—Computers could "learn" from a wide number of sources (decision rules, research data, norms) and "experiences" (modal, as well as novel, clients, feedback on errors and successes) to become increasingly more "intelligent."
9. Each of the *above features might be integrated and interactive*—The test site typically would be different from the site for processing the data.

Technology generally solves some problems while creating others. A recent survey of practitioners indicated considerable ethical concerns related to the use of emerging technologies, such as on-line therapy, direct client services through e-mail, and virtual-reality interventions (McMinn et al., 1999). Unfortunately, as the amount, complexity, interconnectedness, and access to information increases, client confidentiality becomes progressively more of an issue. This seems to be particularly relevant in relation to managed health care which typically requires increasingly more client information to assist with allocating resources. In addition, persons who are adept at various search strategies seem to obtain quite easily detailed information about persons. The above suggests that as client assessment results progressively become more computerized and processed in central locations, potential access to these records also becomes a concern. Additional issues include the continued use of obsolete software, encouragement of a technician role in conducting assessment, temptation to perceive computer-generated narrative reports as finished products, unwarranted faith in the accuracy of computer-based interpretations, use by untrained persons, possible lack of comparability between paper–pencil versus computerized administrations, and potential to de-emphasize the decision-making role of clinicians (Groth-Marnat & Schumaker, 1989; Spielberger & Piotrowski, 1990).

Central to the above concerns is the role of the practitioner in an increasingly automated assessment environment. In the past, one of the crucial roles has been the clinician's responsibility for insuring that interpretations are valid and decisions are appropriate. However, there have been a number of concerns and limitations related to clinician judgement (see Garb, 1998) such that automated interpretations and decisions may be more accurate than ones made by clinicians. This may become increasingly true when data processing involves a large research/normative base, assessment data are highly integrated, the number of variables concerned is extensive, decision trees/processes are complex, and the computer has in-built artificial intelligence. This would call for a change in orientation. The question then becomes how the clinician–computer *interaction* can be optimized to enhance incremental validity. The role for clinicians in the human-oriented process of helping the client to integrate the information also would become more crucial. It also would seem that the clinician would become crucial in monitoring that the process is proceeding optimally. This would be particularly crucial at each of the following transitional phases of assessment: initial referral, intake interview, test administration, integrating feedback, linking the client with additional assessment, extending the results into treatment, or monitoring the progress of interventions.

Visions of Nostradamus

The twenty-first century is likely to bring considerable changes to psychological assessment. Many of these will be extensions of the current and past issues previously outlined. For example, Hunt's (1946) focus on qualitative behavioral observations and clinical judgment always will be important areas on which to focus. Other areas of development include changes in health-care delivery, ecological validity, treatment planning, new theories/tests of abilities, and which healthy aspects of the client to emphasize. As with Hunt's reflections of over 50 years ago, these issues will evolve based on new theories (i.e., intelligence, chaos theory), technologies (i.e., innovations in computer-assisted assessment), changing social values, and an ever-expanding research base.

Patterns of test use over the past thirty years indicate considerable stability (Watkins et al., 1995). During the first 10 to 20 years of the millennium, it is likely that this stability will continue in that the major current tests in the field (Wechsler scales, MMPI-2) will dominate. However, given their current revisions and innovations, they probably have reached their maximum potential and yet there are continual pressures for improvement. Given this, it is difficult to believe that the combination of advances in technology and new theories (even a paradigm shift in the field) would not result in very serious challenges to these major instruments sometime in the next twenty to twenty-five years.

The form of new assessment instruments has been prefaced in the previous section and would be likely to parallel emerging technologies in computer technology, artificial intelligence, and biogenetics. One of the key elements also would be the integration of information to optimize the incremental validity of various sources of information. This process might be assisted by developments in artificial intelligence and a gradually increasing normative database. Normative comparisons thus might be tailored towards such client characteristics as ethnicity, age, gender, and education. Table 1 lists the possible

Table 1
Milestones in the History of Clinical Assessment: 2000–2050

2000	Millennium reflection and evaluation of the field
2003	First convention on computer-assisted assessment strategies involving both professional psychologists and developers of commercially available computer games
2004	Publication of DSM-IV
2005	Virtual reality used to improve clinician judgment by using simulated situations combined with feedback
2008	Publication of WISC-IV
2010	Publication of MMPI-III
2012	Publication of WAIS-IV
2015	Introduction of ability measure that presents major challenge to time-honored scales from the 20th century: Testing situation based on virtual reality, simulation of actual life situations, high ecological validity resulting in high predictions of everyday abilities
2020	New ability measure (2015) supercedes previous ability measures from the 20th century
2025	Knowledge from human genome first used as part of clinical assessment; experimental due to high cost
2035	Widespread use of genome measures for assessment due to increasing cost effectiveness of procedures
2050	First fully integrated assessment instrument using a combination of AI, interactive virtual reality (or possibly hologram), physiological monitors, massive interlinked internet norms, validity/predictions based on chaos theory, branching strategies, genetic measures, in session as well as time series measures

developments in clinical assessment over the first 50 years of the twenty-first century. The culmination of these developments has been listed (conveniently) as 2050 and involves a combination of artificial intelligence, interactive virtual reality, large interlinked norms, physiological monitors, use of mathematical modeling based on chaos theory, multivariate predictive models (such as predicting recidivism in criminal populations), genetic measures, in session as well as measures external to the formal evaluation session (i.e., time-series measures), and the combination of this information to optimize incremental validity. Computers also might be used to assess and interpret narrative information from spoken interactions, written material, or dream content. It is likely that the site of assessment might be quite different from where the analysis and interpretation of the data occur.

One very significant concern is the level of intrusiveness that will be possible and allowable. Currently, if someone wishes to assess another person, they must obtain informed consent and the client must be cooperative. In the future, there may be a number of well-developed techniques for obtaining psychological information without the client even being aware this is occurring. For example, DNA residue from merely touching objects may provide extensive knowledge regarding such clinically relevant information as aspects of personality, intelligence, or the likelihood of developing various types of psychopathology (bipolar disorder, alcoholism). Another strategy might involve computer analysis of narrative content to assess personality, interpersonal style, and predict relevant behaviors. This information might be combined and integrated with unobtrusive physiological measures of tension and response latency through voice and nonverbal analysis. In addition, access to records might become easier given a highly networked, information-oriented society. These essential human-rights issues will need to be given serious consideration, as well as appropriate safeguards.

I have tried to imagine the future of assessment, but I am also humbled by the realization that these changes may be so extensive and unexpected as to be unimaginable. This vision holds considerable promise, as well as numerous challenges. I very much look forward to seeing how they unfold, and especially look forward to both innovation and surprise. It is likely that the developments of the past 50 years will be overshadowed by those over the next 50 years.

References

Acker, M.B. (1990). A review of the ecological validity of neuropsychological tests. In D. Tupper & K.D. Cicerone (Eds.), The neuropsychology of everyday life: Assessment and basic competencies. Boston: Kluwer.

American Psychological Association. (1992). Ethical principles of psychologists and code of conduct. American Psychologist, 47, 1597–1611.

Belter, R.W., & Piotrowski, C. (in press). Training in psychological assessment: A survey of master's degree programs. Journal of Psychological Practice.

Beutler, L.E., Clarkin, J.F., & Bongar, B. (in press). Systematic guidelines for treating the nonbipolar depressed patient. New York: Oxford.

Beutler, L.E., & Williams, O.B. (1998). Systematic treatment selection: A software program. Minneapolis, MN: New Standards, Inc.

Butcher, J.N. (1998). Butcher Psychological Treatment Planning Inventory: Test manual and interpretive guide. San Antonio, TX: Psychological Corporation.

Chen, I.Q., Isberg, E., & Krechevsky, M. (1998). Project Spectrum Early Learning Activities. Cambridge, MA: Harvard Project Zero.

Compass, B.E., Haaga, D.A.F., Keefe, F.J., Leitenberg, H., & Williams, D.A. (1998). Sampling of empirically supported psychological treatments from health psychology: Smoking, chronic pain, cancer, and bulimia nervosa. Journal of Consulting and Clinical Psychology, 66, 89–112.

Davidson, O.R., Stevens, D.E., Goddard, G.V., Bilkey, D.K., & Bishara, S.N. (1987). The performance of a sample of traumatic-head injured patients on some novel computer-assisted neuropsychological tests. Applied Psychology: An International Review, 36, 329–342.

Eisman, E.J., Dies, R.R., Finn, S.E., Eyde, L.D, Kay, G.G., Kubiszyn, T.W., Meyer, G.J., & Moreland, K.L. (1998). Problems and limitations in the use of psychological assessment in contemporary health care delivery: Report of the Board of Professional Affairs Psychological Assessment Work Group, Part II. Washington, DC: American Psychological Association.

Exner, J.E. (1993). The Rorschach: A comprehensive system. Volume I: Basic foundations (3rd ed.). New York: John Wiley & Sons.

Faust, D. (1991). Forensic neuropsychology: The art of practicing a science that does not yet exist. Neuropsychology Review, 2, 205–231.

Finn, S., & Martin, H. (1997). Therapeutic assessment with the MMPI-2 in managed care settings. In J.N. Butcher (Ed.), Personality assessment in managed health care. (pp. 131–152). New York: Oxford University Press.

Fisher, D., Beutler, L.E., & Williams, O. (1999). Making assessment relevant to treatment planning: The STS clinician rating form. Journal of Clinical Psychology, 55, 825, 842.

Fridlund, A., & Delis, D. (1987). CVLT Administration and Scoring System. San Antonio, TX: Psychological Corporation.

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

Gardner, H. (1993). Multiple intelligences: The theory and practice. New York: Basic Books.

Garfield, S.L., & Kurtz, R.M. (1973). Attitudes toward training in diagnostic testing: A survey of directions of internship training. Journal of Consulting and Clinical Psychology, 40, 350–355.

Gass, C.S., & Brown, M.C. (1992). Neuropsychological test feedback to patients with brain dysfunction. Psychological Assessment, 4, 272–277.

Golden, C., Purisch, A.D., & Hammeke, T.A. (1985). Luria–Nebraska Neuropsychological Battery: Forms I and II. Los Angeles: Western Psychological Services.

Griffith, L.F. (1997). Surviving no-frills mental healthcare: The future of psychological assessment. Journal of Practical Psychiatry and Behavioral Health, 3, 255–258.

Groth-Marnat, G. (in press). Financial efficacy of clinical assessment: Rational guidelines and issues for future research. Journal of Clinical Psychology.

Groth-Marnat, G. (Ed.). (in press). Neuropsychological assessment in clinical practice: A practical guide to test interpretation and integration. New York: John Wiley & Sons.

Groth-Marnat, G. (1999). Handbook of psychological assessment (3rd ed., rev.). New York: John Wiley & Sons.

Groth-Marnat, G., Gallagher, R.E., Hale, J.B., & Kaplan, E. (2000). The Wechsler intelligence scales in neuropsychological assessment. In G. Groth-Marnat (Ed.), Neuropsychological assessment in clinical practice: A practical guide to test interpretation and integration (pp. 129–194). New York: John Wiley & Sons.

Groth-Marnat, G., & Schumaker, J. (1989). Computer-based test interpretation: Issues and guidelines. American Journal of Orthopsychiatry, 59, 257–263.

Groth-Marnat, G., & Teal, M. (in press). Block design as a measure of everyday spatial ability: An ecological validity study. Perceptual and Motor Skills.

Hammainen, L. (1994). Computerized support for neuropsychological test interpretation in clinical situations. The Clinical Neuropsychologist, 8, 167–185.

Harkness, A.R., & Lilienfeld, S.O. (1997). Individual differences science for treatment planning. Psychological Assessment, 9, 349–360.

Harrison, P.L., Flanagan, D.P., & Genshaft, J.L. (1997). An integration and synthesis of contemporary theories, tests, and issues in the field of intellectual assessment. In D.P. Flanagan, J.L. Genshaft, & P.L. Harrison (Eds.), Contemporary intellectual assessment: Theories, tests, and issues. New York: Guilford.

Harrison, P.L., Kaufman, A.S., Hickman, J.A., & Kaufman, N.L. (1988). A survey of tests used for adult assessment. Journal of Psychoeducational Assessment, 6, 188–198.

Heiby, E.M. (1995). Chaos theory, nonlinear dynamical models, and psychological assessment. Psychological Assessment, 7, 5–9.

Hunt, W.A. (1946). The future of diagnostic testing in clinical psychology. Journal of Clinical Psychology, 2, 311–317.

Jacobson, N.S., Follette, W.C., & Pagel, M. (1986). Predicting who will benefit from behavioral marital therapy. Journal of Consulting and Clinical Psychology, 54, 518–522.

Kane, R.L., & Kay, G.G. (1992). Computerized assessment in neuropsychology: A review of tests and test batteries. Neuropsychology Review, 3, 1–117.

Kaplan, E., Fein, D., Morris, R., & Delis, D. (1991). WAIS-R as a neuropsychological instrument. San Antonio, TX: The Psychological Corporation.

Kaufman, A.S., & Kaufman, N.L. (1983). Interpretive manual for the Kaufman Assessment Battery for Children. Circle Pines, MN: American Guidance Service.

Kinder, B.N. (1994). Where the action is in personality assessment. Journal of Personality Assessment, 62, 585–588.

Larrabee, G.J., & Crook, T. (1988). A computerized everyday memory battery for assessing treatment effects. Psychopharmacology Bulletin, 24, 695–697.

Lubin, B., Larsen, R.M., & Matarazzo, J.D. (1984). Patterns of psychological test usage in the United States: 1935–1982. American Psychologist, 39, 451–454.

Lubin, B., Larsen, R.M., Matarazzo, J.D., & Seever, M. (1985). Psychological test usage patterns in five professional settings. American Psychologist, 40, 857–861.

Lubin, B., Larsen, R.M., Matarazzo, J.D., & Seever, M. (1986). Selected characteristics of psychologists and psychological assessment in five settings: 1959–1988. Professional Psychology: Research and Practice, 17, 155–157.

McMinn, M.R., Buchanan, T., Ellens, B.M., & Ryan, M.K. (1999). Technology, professional practice, and ethics: Survey of findings and implications. Professional Psychology: Research and Implications, 30, 165–172.

Meehl, P.E. (1954). Clinical versus statistical prediction: A theoretical analysis and a review of the evidence. Minneapolis, MN: University of Minnesota Press.

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 1. Washington, DC: American Psychological Association.

Naglieri, J.A. (1999). Essentials of CAS assessment. New York: John Wiley & Sons.

Parker, K.C. (1983). A meta-analysis of the reliability and validity of the Rorschach. Journal of Personality Assessment, 47, 227–231.

Parker, K.C., Hanson, R.K., & Hunsley, J. (1988). MMPI, Rorschach, and WAIS: A meta-analytic comparison of reliability, stability, and validity. Psychological Bulletin, 103, 367–373.

Piotrowski, C. (1999). Assessment practices in the era of managed care: Current status and future directions. Journal of Clinical Psychology, 55, 787–796.

Piotrowski, C., & Belter, R.W. (1999). Internship training in psychological assessment: Has managed care had an impact? Manuscript submitted for publication.

Piotrowski, C., & Zalewski, C. (1993). Training in psychodiagnostic testing in APA-approved PsyD and PhD clinical psychology programs. Journal of Personality Assessment, 61, 394–405.

Pope, K.S. (1992). Responsibilities in providing psychological test feedback to clients. Psychological Assessment, 4, 268–271.

Prochaska, J.O., DiClemente, C.C., & Norcross, J.C. (1992). In search of how people change: Applications to addictive behaviors. American Psychologist, 47, 1102–1114.

Psychological Corporation. (1998). OPTAIO [Computer software]. San Antonio, TX: The Psychological Corporation.

Rapaport, C., Gill, M., & Schafer, J. (1946). Diagnostic psychological testing (Vol. 2). Chicago: Year Book Publishers.

Russell, E.W. (1998). In defense of the Halstead Reitan Battery: A critique of Lezak's review. Archives of Clinical Neuropsychology, 13, 365–381.

Sawyer, J. (1966). Measurement and prediction, clinical and statistical. Psychological Bulletin, 66, 178–200.

Sbordone, R.J., & Long, C.J. (Eds.). (1996). Ecological validity of neuropsychological testing. Delray Beach, FL: GR Press/St. Lucie Press.

Spielberger, C.D., & Piotrowski, C. (1990). Clinician's attitudes toward computer-based testing. The Clinical Psychologist, 43, 60–63.

Sternberg, R., Wagner, R.K., Williams, W.M., & Horvath, J.A. (1995). Testing common sense. American Psychologist, 50, 912–927.

Stout, C.E., & Cook, L. (1999). New areas for psychological assessment in general health care settings: What to do today to prepare for tomorrow. Journal of Clinical Psychology, 55, 797–812.

Strohmer, D.C., & Newman, L.J. (1983). Counselor hypothesis testing strategies. Journal of Counseling Psychology, 30, 557–565.

Sundberg, N. (1961). The practice of psychological testing in clinical services in the United States. American Psychologist, 16, 79–83.

Sweet, J.J., Moberg, P.J., & Westergaard, C.K. (1996). Five-year follow-up survey of practices and beliefs of clinical neuropsychologists. The Clinical Neuropsychologist, 10, 202–221.

Task Force on Promotion and Dissemination of Psychological Procedures. (1995). Training in and dissemination of empirically-validated psychological treatments: Report and recommendations. The Clinical Psychologist, 48, 3–23.

Wagner, R.K., & Sternberg, R.J. (1991). Tacit knowledge for managers. San Antonio, TX: Psychological Corporation.

Watkins, C.E., Campbell, V.L., Nieberding, R., & Hallmark, R. (1995). Contemporary practice of psychological assessment by clinical psychologists. Professional Psychology: Research and Practice, 26, 54–60.

Wedding, D. & Faust, D. (1989). Clinical judgment and decision making in neuropsychology. Archives of Clinical Neuropsychology, 4, 233–265.

Weiner, I.B. (1983). The future of psychodiagnostics revisited. Journal of Personality Assessment, 47, 451–459.

Weiss, D.J. (1985). Adaptive testing by computer. Journal of Consulting and Clinical Psychology, 53, 774–789.

Wickramasekera, I. (1995). Concepts, data, and predictions from the high risk model of threat perception. Journal of Nervous and Mental Disease, 183, 15–23.

Wilson, B.A., Cockburn, J., & Baddeley, A. (1985). The Rivermead Behavioral Memory Test. Gaylord, MI: National Rehabilitation Services.

Woodcock, R.W., & Johnson, M.B. (1989). Woodcock–Johnson Tests of Cognitive Ability: Standard and supplemental batteries. Allen, TX: DLM/Teaching Resources.