

Jack A. Cummings
Indiana University

Marc C. Laquerre
Indiana University

Psychoeducational Assessment Research in the United States: From sociocultural context to extrapolations into the future

Abstract

Sociocultural factors are presented in order to provide a frame from which to view psychoeducational assessment research in the domains of cognitive ability, academic achievement, adaptive behavior, and social-emotional functioning. The role of theory is examined as it has had an impact on development and revision of recently published intelligence measures. Research on the complex issue of test bias is also briefly addressed. Two research trends in achievement are highlighted: the mismatch of local curriculum and nationally norm-referenced test content and curriculum-based assessment. The significant developments in the areas of adaptive behavior, social skills, and social-emotional assessment are explored. The paper is concluded with the authors' impressions and recommendations gleaned from the current analysis of research articles, test reviews, and technical manuals.

Résumé

Les facteurs socio-culturels servent de cadre à l'analyse des recherches sur l'évaluation psychopédagogique dans les domaines des compétences cognitives, des résultats scolaires, du comportement adaptatif, du fonctionnement socio-affectif, etc. Les auteurs analysent le rôle de la théorie et son impact sur la conception et le remaniement de mesures récemment publiées sur l'intelligence. Ils abordent également la question complexe des tests biaisés. Deux tendances se dégagent plus particulièrement: la disparité des programmes locaux et du contenu des tests nationaux et l'évaluation basée sur les programmes. Les auteurs analysent les progrès importants réalisés dans les domaines du comportement adaptatif, des habiletés sociales et de l'évaluation socio-affective. L'article s'achève sur les impressions et recommandations des auteurs qui sont le fruit de leur analyse d'articles de recherche, de tests et de manuels techniques.

Research and practice in psychoeducational assessment take place within a sociocultural context. Legislation, litigation over discriminatory aspects of testing, and public opinion have been powerful forces which have shaped psychoeducational assessment with school-age children. The intent of the first section of this paper is to provide a partial sociocultural perspective from which to view psychoeducational research. In subsequent sections two domains of psychoeducational assessment will be examined. These sections will focus on the research which addresses the use of intelligence, academic achievement, adaptive behavior, and social-emotional functioning measures. Each section will begin with a general discussion of the domain. Also included will be an analysis of present practices with emphasis placed on individual psychoeducational assessment to identify and remediate students who experience significant adjustment difficulties in the regular classroom. This will be followed by an examination of research trends within each area. A section on the authors' impressions of research across the three assessment domains will conclude the paper.

Sociocultural Context

In 1975 PL 94-142 (Education for All Handicapped Children Act) was passed by the U.S. Congress, with an effective date of September 1, 1978. It guaranteed all handicapped children the right to a free appropriate education and a nondiscriminatory evaluation. This legislation has institutionalized the assessment of large numbers of children (Salvia & Ysseldyke, 1988). If a child is referred by a teacher or parent, he or she has the right to a nondiscriminatory evaluation. PL 94-142 represents a categorical approach to special education. Definitions are given for deaf, mentally retarded, multihandicapped, orthopedically impaired, seriously emotionally disturbed, and learning-disabled categories of children. An evaluation is conducted to determine if a child is eligible for placement in the area of the suspected handicapped. If the child is judged by a multidisciplinary team to be eligible, he or she is placed in the qualifying category of a special education class.

As noted by Will (1986), PL-142 has had a tremendous impact on the provision of services to children who have difficulty coping with the demands of the regular class environment. The legislation led to the institutionalization of individualized instruction, increased the role of the parents in the education of their handicapped child, and made education possible for 500,000 previously unserved seriously handicapped children.

The Regular Education Initiative (REI) is a movement that suggests regular, not special, educators should assume responsibility for the education of students we would currently classify as handicapped (Reynolds, Wang, &

Walberg, 1987; Stainback & Stainback, 1984; Will, 1986). REI is a reaction to frustrations associated with the current dual system, where regular is segregated from special education. Reynolds *et al.* (1987) are critical of the existing approach in which programs are categorical, e.g., learning-disabled, mildly mentally handicapped, and seriously emotionally disturbed. The problem is that children with significant needs often do not fit a special education category. The debate of the appropriateness of this radical restructuring of special education is prominent in the special education and school psychology journals (Braaten, Kauffman, Braaten, Polsgrove, Nelson, 1988; Gerber, 1988; Keogh, 1988). Additionally, discussions on the merits of the approach are taking place at state and national conferences. The impact of the movement has yet to be realized, because the primary discussants are individuals associated with universities, rather than school-based practitioners (Davis, 1988). REI represents more than a restructuring of special education, because in order to succeed significant reforms will have to be made by regular educators. Liberman (1985) aptly points out that special educators will have a difficult time obtaining support for the regular education initiative from regular educators. While the movement has yet to have a pervasive impact on services, it does reflect a growing criticism of special education.

The assessment of intellectual functioning of minority children has been the focus of a storm of criticism. Plaintiffs in the Larry P. (1979) and PASE (1980) cases contended that intelligence tests were discriminatory and the major cause for disproportionate representation of minorities in special education classes. The federal judge in the former case banned the use of intelligence tests with minorities. The opinion of the court in the PASE case directly contradicted the ruling in Larry P. The conflicting rulings by two circuit courts increases the likelihood that the U.S. Supreme Court will hear the issue of intelligence tests and their alleged discriminatory impact.

Assessment of Intellectual Functioning

The sociocultural forces influencing psychoeducational assessment are clearly evident in the domain of intelligence. Recent landmark litigation and legislation have had a noticeable impact on current intelligence assessment and, undoubtedly, will continue to shape its future. However, not only is the practice of psychoeducational assessment affected by sociocultural forces, the construct of intelligence itself is conceptualized within a socio-cultural context. Sternberg and Salter (1985) have argued that, ". . . all intelligent behavior occurs in a social context that includes goals, expectations, demands and a history of prior experiences" (p. 15). They continue stating that intelligence is defined as "goal-directed, adaptive behavior."

Within the last century, numerous theories have been proffered attempting to describe and explain intelligence. In fact, at times it seems as if there are as many definitions of intelligence as there are theorists. It was this state of affairs that prompted Sternberg to state, "Intelligence is among the most elusive of concepts. Certainly, there are few other concepts that have been conceptualized in as many different ways" (Sternberg, 1985a, p. 3).

Mayr (1982) suggested a dimension which distinguishes two broad categories of intelligence theories. Mayr proposed distinguishing theorists as either "lumpers" or "splitters." Theorists falling in the lumpers category are those who conceive of intelligence as a "general, unified capacity." Spearman's (1927) *g*-factor theory is an early example of this unitary capacity. Likewise, Wechsler's definition of intelligence proposes that intelligence is "the aggregate or global capacity of the individual to act purposefully, to think rationally, and to deal effectively with his environment" (Wechsler, 1944, p. 3). In contrast others suggest that intelligence is not a unitary construct, but rather, composed of a number of separate abilities that are fairly independent of each other, e.g., Guilford (1967). Most theories of intelligence have emerged from a psychometric approach. The structures of intelligence were the primary foci of these theories. In recent years, advances in cognitive psychology have offered new perspectives on the nature of human intelligence (Curtis & Glaser, 1984; Sternberg, 1986; Weinberg, 1989). Sternberg (1985b) suggests that these two approaches be viewed as complementary, with cognitive theories providing welcome insights regarding the cognitive processes that underlie intelligence and intelligent behaviors. Increases in scientific interest and research in information processing has caused the focus to shift from structure to the problem-solving process.

The *Kaufman Assessment Battery for Children* (K-ABC) (Kaufman & Kaufman, 1983) provides an illusion of an intelligence test where the authors emphasized process rather than content. Using Das, Kirby, and Jarman's (1975; 1970) successive and simultaneous model of information processing, Kaufman and Kaufman (1977) set out to build a measure which had as its base a grounding in theory. This stood in stark contrast to Binet's and Wechsler's approach to test development, both of which were atheoretical.

The success with which the K-ABC authors successfully operationalized a processing model has been questioned. Sternberg (1984) pointed out that a selected set of subtests is proposed to measure simultaneous processing, while another set is intended to assess sequential processing. Process is equated with tasks. If the focus had truly been on process of problem-solving, then the test authors would have assessed whether an individual worked at a given task in a successive or simultaneous fashion (Sternberg, 1984).

The role of theory has continued to play a part in the development/revision of cognitive tests during the latter half of the 1980s. As noted by Woodcock (1989), Horn and Cattell's *Gf - Gc* theory (Cattell, 1963; Horn, 1985, 1986) led to a reconstruction of the *Woodcock-Johnson Psycho-Educational Battery-Revised* (Woodcock & Johnson, 1989). In the fourth edition of the Stanford-Binet, Thorndike, Hagen, and Sattler (1986) used a three-level hierarchical model of cognitive abilities to guide the construction of the scale. See Figure 1 for a graphic representation of the theoretical model upon which the Revised Stanford-Binet was built.

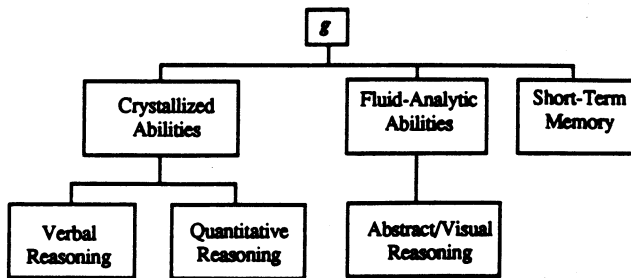


Figure 1
Theoretical Model of the Stanford-Binet: Fourth Edition

Reprinted with permission of The Riverside Publishing Company from page 4 of *Stanford-Binet Intelligence Scale Guide for Administering and Scoring the Fourth Edition*, by R.L. Thorndike, E.P. Hagen, and J.M. Sattler. The Riverside Publishing Company, 8420 W. Bryn Mawr Avenue, Chicago, IL 60631. Copyright 1986.

While recent test revisions have reflected major shifts in thinking about the theoretical construct of intelligence, the actual practice of assessment has shown signs of lagging behind. Federal and state definitions of mental retardation and learning disabilities reflect the unitary concept of intelligence. Likewise, an unfortunate tendency for practitioners has been to judge new measures against the past. Based on the assessment of hundreds or thousands of children, practitioners develop internal norms. When a new measure becomes available it is unfairly judged against the tried and true.

While acknowledging the importance of appropriate and accurate instrumentation, one must not diminish the importance of the psychologist in intelligence testing. A critical link in the chain of intelligence assessment is the examiner. Test scores remain to be interpreted. In elaborating the concept of "intelligent testing," Kaufman (1979) emphasizes the crucial role of the psychologist. He presents a well-founded and logical argument that test scores be interpreted within a more general context that includes observa-

tions, and cognitive and developmental theories. Such interpretation requires knowledge, flexibility, and insight on the part of the examiner. Technically adequate intelligence tests and well-prepared, observant psychologists are partners in the assessment enterprise and must not be seen as substitutes for each other.

In the 1970s and early 1980s, issues concerning the discriminatory nature of intelligence assessment measures received much public and professional attention. Measures were subjected to intensive scrutiny for evidence of test bias (Kaplan, 1985). Test bias is not a novel concept. It has received intermittent scholarly attention since the nascence of intelligence testing. In recent years, however, sparked by the controversial decisions of Larry P. (1979) and PASE (1980), the issue of test bias has been subject to renewed scientific inquiry and research.

At the heart of this controversy is the question of whether contemporary intelligence tests are biased against certain ethnic and cultural minorities. That certain minority groups typically score lower on tests of intellectual ability has been demonstrated for some time (see Jensen, 1980, for a review). This fact alone has, at times, been cited as evidence of test bias. It has also prompted emotional outbursts condemning intelligence tests as agents for the perpetuation of racism (Jackson, 1975; Williams, 1974). Unfortunately, these outbursts may have done more to obscure the issues than to clarify. Reynolds and Brown (1984) accurately characterize the question of test bias as an empirical one. As such, it demands evaluation within the rigors of scientific methodology.

Some have suggested that there is no *a priori* explanation for the existence of differences in performance on tests of intellectual ability across race (Alley & Foster, 1978; Hilliard, 1979). Reynolds and Brown (1984) replied that there is, likewise, no *a priori* basis for believing that differences should not exist. They state that, "Group differences on mental tests, or any other kind of test, *per se* give no directly applicable information regarding test bias" (Reynolds & Brown, 1984, p. 24). Jensen (1980) has suggested that the argument that simple group mean differences is evidence of test bias assumes the egalitarian fallacy (i.e., that all human subgroups are identical or equal in every trait) and reflects an unsophisticated understanding of test bias.

Distinctions have been made between test bias and test misuse (Reynolds, 1982; Reynolds & Brown, 1984). Test misuse concerns the use of a test or test results to make decisions that are unfair or biased. Test misuse can affect individuals from any group. Test bias, however, is typically defined in purely psychometric terms. Reynolds and Brown (1984) define test bias as "a

statistical term referring to a constant error of a measure in one specific direction as opposed to random error" (p. 2). Given this definition, test bias is essentially an issue of validity (Cole, 1981; Jensen, 1980; Reynolds, 1982). Numerous researchers have considered the implications and evaluation of test bias in the domains of content, construct, and predictive validity.

Another current in recent research of test bias is based on the Spearman hypothesis. Jensen (1980) has been a leading, contemporary proponent of this hypothesis. Generally, the Spearman hypothesis suggests that black-white differences in performance on intelligence measures are attributable to differences in *g*-factor loading. The magnitude of black-white differences is hypothesized to vary directly with *g*-factor loadings. Some subsequent research supported this hypothesis (Jensen, 1985). Tests with high *g*-loading yield greater black-white difference in performance. In a more recent study, Naglieri and Jensen (1987) compared the black-white differences in performance on the *Wechsler Intelligence Scale for Children – Revised* (WISC-R) and on the K-ABC. Their findings supported the Spearman hypothesis and indicated that the lower black-white difference demonstrated on the K-ABC was due, in part, to lower *g*-loadings. Jensen's work and the Spearman hypothesis, however, are not without critics. Humphreys (1985) explored black-white differences controlling for race and socioeconomic status. The results of his study suggested that the Spearman hypothesis more accurately applied to differences in socioeconomic status than to differences in race.

The controversy of bias has demonstrated that issues of test bias and such corollary concerns as racial discrimination and societal inequality can become conceptually entwined. Hopefully, a beneficial outcome of the intense scrutiny of test bias has been the development of a better and more accurate understanding. Certainly, recent test developers appear more sensitive to the issue (Kaufman & Kaufman, 1983). Cole (1981) concluded that while scientists can examine such technical aspects of psychometric tests as test bias, they cannot answer the more vexing questions of social policy.

Assessment of Academic Achievement

The focus of achievement testing is to assess a student's attainment of academic skills. These often include reading, written language, and mathematical functioning. Anastasi (1982) notes that, traditionally, academic assessment is distinguished from cognitive/intellectual assessment by the degree to which the measure is designed to assess uniform versus diverse antecedent experiences. Since first grade reading instruction is a fairly uniform previous experience, a test designed to measure reading attainment would be categorized as an academic achievement measure. In contrast, an intellectual assess-

ment measure attempts to measure diverse antecedent experiences, e.g., verbal reasoning, spatial skills, memory, etc. As Cummings (1987) states, "contemporary measurement specialists recognize that both achievement and aptitude tests assess acquired knowledge, but differ on the degree of specificity and abstraction" (p. 21).

A significant development in the assessment of academic achievement is the tendency of test publishers to promote batteries which include both cognitive and achievement measures normed on the same standardization sample. The *Kaufman Assessment Battery for Children* (Kaufman & Kaufman, 1983) and the revised Woodcock-Johnson exemplify this trend. Both of these batteries include cognitive sections to assess intellectual functioning as well as academic achievement.

There has been considerable controversy over the Kaufmans' decision to include what some have labeled measures of verbal intelligence on the achievement scale (Anastasi, 1984; Jensen, 1984; Keith, 1985, 1986; Keith & Dunbar, 1984; Sternberg, 1984). Placing the more cognitively complex subtests (Expressive Vocabulary, Faces & Places, and Riddles) on the achievement scale led to reduced black-white differences, but also removed highly g-loaded measures from the estimate of intelligence provided by the mental processing composite (a combined score from the simultaneous and successive processing scales).

A significant trend in research has been the examination of the mismatch of curriculum and norm-referenced test content when interpreting students' performances on norm-referenced measures (Good & Salvia, 1988; Jenkins & Pany, 1978; Leinhardt & Seewald, 1981; Mehrens, 1984; Mehrens & Phillips, 1986; Schmidt, 1983). At issue is the lack of overlap between the curriculum and the test. When a student's test performance is systematically influenced by a difference between the content of the curriculum and the test, it is referred to as curriculum (or content) bias. A study by Good and Salvia (1988) will provide an example. They administered four reading-achievement tests to 65 students who were exposed to a specific reading curriculum. They did a word analysis to quantify the overlap between the four reading measures and the reading program. As expected the students' reading-test performance could be predicted based on the degree of overlap between the test and the curriculum to which they were exposed.

Curriculum-based assessment (also referred to as curriculum-based measurement) has been proposed as an alternative to the norm-referenced approach embodied in both federal and state special education legislation. Curriculum-based assessment focuses on the curricula to which the student is

exposed. This reduces the mismatch between the student's own instructional experiences and the curriculum reflected in the selection of norm-referenced test items. Deno and his colleagues (Deno, 1985, 1986; Jenkins, Deno, & Mirkin, 1979; Marston, Lowry, Deno-Mirkin, 1981; Shinn & Marston, 1985) have described procedures for assessing student achievement in reading, spelling, and arithmetic. For instance, to assess a student's reading skills, he or she would be asked to read aloud for one minute, with the examiner recording the number of words read correctly and incorrectly. This type of reading sample would be repeated frequently, e.g., daily, so as to yield time-series data that could be used to judge the impact of a given instructional program (Deno, 1986). Another example of curriculum-based assessment, applied to written language, would be to use a story starter and allow a student to write for three minutes. Scoring can be as simple as counting correctly written word or letter sequences, or counting more complex, low frequency words.

Adaptive Behavior and Social-Emotional Assessment

Within the last 15 years significant strides have been taken in the domain of adaptive behavior assessment. Prior to the publication of Mercer and Lewis's (1977) *Adaptive Behavior Inventory for Children*, the best adaptive behavior measures failed to meet even minimal psychometric criteria. "Clinical judgement" was the best available practice for evaluating a child's adaptive behavior functioning. Despite the recognition of the importance of adaptive behavior in the legislation covering the identification of mildly mentally handicapped children, the IQ score often was the sole determining factor, with adaptive behavior being ignored. Since the publication of Mercer's adaptive behavior measure, the revised *Vineland Adaptive Behavior Scales* (Sparrow, Balla, & Cicchetti, 1984) and the *Scales of Independent Behavior* (Bruininks, Woodcock, Weatherman, & Hill, 1984) have been published. Both of these reflect careful attention to the test development process and have adequate nationally representative standardization samples (Cummings & Simon, 1988).

Harrison (1989) is correct in her observation that the bulk of research on adaptive behavior has focused on the psychometric properties of the scales. Despite this psychometric emphasis, Harrison (1989) notes that insights have been gained into the multifaceted nature of adaptive behavior through concurrent validation studies. As expected the communication domain of the adaptive behavior (as opposed to the self-help, socialization, or domestic skills domains) correlates most highly with intelligence (Bruininks & McGrew, 1987; Harrison & Kamphaus, 1984; Keith, Fehrmann, Harrison, & Pottebaum, 1987).

Several researchers have investigated the relationship between parents' and teachers' judgments of a child's adaptive behavior functioning (Heath & Obrzut, 1986; Meador & Richmond, 1980; Wall & Paradise, 1981). These researchers have reported that parents' ratings of their children's adaptive behavior are significantly higher than those scores obtained when teachers served as informants.

Depending on one's perspective, social skills may be subsumed as one domain of adaptive behavior functioning or, as Gresham and Reschly (1988) propose, both adaptive behavior and social skills are components of social competence. Although Gresham and Reschly's model is useful as a framework for a practitioner's conceptualization of an assessment, a unified theoretical approach to adaptive behavior/social skills has yet to be advanced and supported with empirical evidence. Notwithstanding, the technical sophistication of behavioral approaches to the assessment of social skills has improved dramatically due to the work of Elliott, Sheridan, and Gresham (1989), Gresham and Elliott (1984, 1989), and Gresham and Reschly (1988). Gresham and Reschly (1988) present systematic procedures for understanding social skill deficits, performance deficits, and self-control deficits. A strength of their work has been the emphasis on treatment validity. Assessment is judged by the degree to which it provides information relevant to the development and implementation of an intervention. Witt and Elliott (1986) have conducted an impressive array of treatment-acceptability studies, examining attitudes of those who assist with or have primary responsibility for implementation of an intervention.

Gresham and Elliott (1989) recently published the *Social Skills Rating System*. It is designed to capture behaviors that affect the teacher-student relationship, academic performance, and peer acceptance. There are three forms, one each for collecting the impressions of teachers, parents, and students. The system was constructed to ensure the social validity of behaviors selected as targets for intervention.

A measure in the social-emotional domain which has improved significantly is the *Revised Children's Manifest Anxiety Scale* (RCMAS) (Reynolds & Richmond, 1985). It has benefited from the efforts of several researchers. Reynolds and Paget (1983) reported the findings of a national standardization effort that included 4,972 children from age 6 to 19. Saigh (1989) and Mattison, Bagnato, and Brubaker (1988) examined and found merit in the use of the RCMAS with the DSM-III. Whereas Reynolds and Richmond's (1978) first attempt at revising the *Children's Manifest Anxiety Scale*, which they called "What I think and feel," would have been criticized for lacking a manual and adequate norms, the RCMAS was endorsed by Gresham (1989)

in the most recent *Mental Measurements Yearbook*. The RCMAS is illustrative of other scales which have profited from an investment of time and effort associated with national standardization samples, and the collective efforts of independent researchers conducting reliability and validity studies.

The *Martin Temperament Assessment Battery for Children* (Martin, 1988) was designed as a measure of temperament variables, activity level, adaptability to new environments, approach/withdrawal tendencies, emotional intensity, distractibility, and persistence. In addition to using the parent and teacher as sources of information, Martin also devised a scale for the clinician to rate a child's temperament during the evaluation process. The standardization sample for the teacher and parent forms was drawn from three regions of the country (Southeast, Northeast, and Rocky Mountain regions). Thus it is not representative of the country as a whole and the scale should be viewed as being in its early phase of development.

Whereas significant developments have occurred in the adaptive-behavior and rating-scale approaches to social-emotional functioning, our view of the progress made in the personality assessment domain, i.e., projective testing, is less favourable.

The *Personality Inventory for Children* (PIC) has been the object of a significant amount of research. Most of the studies have involved the comparison of a clinically identified sample with "normal" peers or other clinically different samples (cf. Clark, Kehle, & Bullock, 1988; Ehly, Reimers, & Keith, 1986; Kelly, 1988). Despite a growing research base on the PIC, its popularity among clinicians, and the publication of revised test manuals (Lachar, 1982; Wirt, Lachar, Klinedinst, & Seat, 1984), the use of the PIC is inappropriate due to the absence of a recent and nationally representative standardization sample. Reviewers for the *Mental Measurements Yearbooks* have harshly criticized the publishers of the PIC for failing to mount a national norming effort (Knoff, 1989; Reynolds, 1985).

Projective drawing techniques, although widely used, continue to be plagued by basic validity problems. Too often the research that is conducted on children's drawings compares the drawing of a clinically identified group with normal controls. The findings of these studies are often contradictory (Cummings, 1986). (For a review of thematic approaches the reader is referred to Obrzut and Cummings [1983].)

Conclusions and Recommendations for the Future

Trying to capture the essence of recent research in the area of psychoeducational assessment is a difficult proposition. It is akin to the fabled doctoral qualifying question of defining the universe and listing three ex-

amples. It would be an understatement to say that researchers are pursuing many different paths. This paper provides coverage of a handful of these directions. Many others should have been included in the review, but had to be omitted due to the space limitations. Among those researchers whose work has considerable merit but has been omitted are: Naglieri (1985), *Matrix Analogies Test*; Achenbach and McConaughy (1987), *Multiaxial Empirically Based Assessment*; and Ysseldyke and Christenson (1987), *The Instructional Environments Scale*.

From our review of a large number of research articles, test reviews, and test manuals we offer the following impressions and recommendations. A substantial amount of research is being conducted under the guidance of test publishers. This has both positive and negative facets. On the positive side, the research conducted by test publishers has meant that much is known about the psychometric properties of a measure at the time of publication. More complete technical manuals mean that consumers are better able to judge the adequacy of the test. There is no question that the vast quantities of applied assessment research have resulted in measures that are better developed and standardized. However, a negative aspect is that much of this research is being approved, directed, and often interpreted by the test author and/or publisher. Clearly, due to the financial investment of the publisher, there is potential for bias. As a consequence, researchers need to have access to publishers' data, in addition to pursuing comprehensive independent investigations.

Keith (1987) makes a good point when he suggests that researchers conducting investigations with tests need to approach the enterprise with a hypothesis-testing perspective. All too often, researchers merely correlate two measures, "simple minded correlations of one test with another for no apparent reason beyond publication" (Keith, 1987, p. 276). There are signs that the sophistication of assessment researchers is improving. The increased use of theory is encouraging. The role that theory has played in the development/revision of the three most recently published cognitive measures (*Kaufman Assessment Battery for Children*, *Stanford-Binet IV*, *Woodcock-Johnson Tests of Cognitive Ability-Revised*) indicates that progress is being made. Theory derived from advances in cognitive science should inform our future hypothesis testing efforts to develop and interpret psychoeducational measures.

The question of why research is frequently ignored by practitioners is a genuine enigma. Empirical investigations have consistently failed to support the use of projective drawings, yet the use of drawings in psychological

batteries is commonplace. Likewise, misinterpretation of WISC-R scatter is rampant and is part of the clinical folklore of too many multidisciplinary teams, yet Kaufman's (1976) analysis of the standardization data and "normal" scatter was published more than a decade and a half ago.

The potential for the use of computers as assessment tools within educational environments has only partially been realized. The availability of powerful microcomputers in combination with item-response theory means that sophisticated adaptive testing procedures have the potential of revolutionizing the way assessment is conducted. Weiss (1982, 1983; Weiss & Kingsbury, 1984) demonstrated that computerized adaptive testing procedures have reliabilities and validities equal to or better than those of comparable conventional measures. By incorporating the use of item-response theory, the computer estimates trait levels and administers items based on the individual's successful or unsuccessful responses to previous items. Thus the number of items administered can be reduced by as much as 50 percent by tailoring the test to the individual. Interactive video enhances the potential of assessment by adding real-time audio and video images. Previously videotaped instructions or situations are stored on a videodisc and may be recalled almost instantaneously by the press of a key.

In conclusion, our understanding of the constructs we attempt to measure is clearer and test developers/publishers are beginning to recognize the importance of psychometric standards. Our challenge is to conduct research that allows us to view the individual we assess from a more holistic perspective, not one fragmented by a litany of separate tests.

REFERENCES

- Achenbach, T.M., & McConaughy, S.H. (1987). *Empirically-based assessment of child and adolescent psychopathology: Practical applications*. Newbury Park, CA: Sage.
- Alley, G., & Foster, C. (1978). Nondiscriminatory testing of minority and exceptional children. *Focus of Exceptional Children*, 9, 1-14.
- Anastasi, A. (1982). *Psychological Testing* (Fifth Ed.). New York: Macmillan.
- Anastasi, A. (1984). The K-ABC in historical and contemporary perspective. *Journal of Special Education*, 18, 357-366.
- Braaten, S., Kauffman, J.M., Braaten, B., Polsgrove, L., & Nelson, C.M. (1988). The regular education initiative: Patent medicine for behavior disorders. *Exceptional Children*, 55, 21-27.

- Bruininks, R.H., & McGrew, K. (1987). *Exploring the structure of adaptive behavior (Report Number 87-1)*. Minneapolis: University of Minnesota, Department of Educational Psychology.
- Bruininks, R.H., Woodcock, R.W., Weatherman, R.F., & Hill, B.K. (1984). *Scales of Independent Behavior: Woodcock-Johnson Psycho-Educational Battery - Part Four*. Allen, TX: DLM Teaching Resources.
- Cattell, R.B. (1983). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology, 54*, 1-22.
- Clark, E., Kehle, T.J., & Bullock, D.S. (1988). Personality inventory for children: Profiles for learning disabled, emotionally disabled, and intellectually handicapped children. *School Psychology International, 9*, 43-49.
- Cole, N.S. (1981). Bias in testing. *American Psychologist, 36*, 1067-1077.
- Cummings, J.A. (1987). Academic assessment. In C.R. Reynolds & L. Mann (Eds.), *Encyclopedia of Special Education*, Vol. 1 (pp. 20-24). New York: John Wiley.
- Cummings, J.A. (1986). Projective drawings. In H.M. Knoff (Ed.), *The psychological assessment of child and adolescent personality* (pp. 199-244). New York: Guilford Press.
- Cummings, J.A., & Simon, M.S. (1988). Review of the Scales of Independent Behavior. *Journal of Psychoeducational Assessment, 6*, 315-320.
- Curtis, M.E., & Glaser, R. (1984). Intelligence testing, cognition and instruction. *International Journal of Psychology, 19*, 475-497.
- Das, J.P., Kirby, J.R., & Jarman, R.F. (1975). Simultaneous and successive syntheses: An alternative model for cognitive abilities. *Psychological Bulletin, 82*, 87-103.
- Das, J.P., Kirby, J.R., & Jarman, R.F. (1979). *Simultaneous and Successive Cognitive Processes*. New York: Academic Press.
- Davis, W.E. (1988). The regular education initiative debate: Its promises and problems. *Exceptional Children, 55*, 40-446.
- Deno, S.L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children, 52*, 219-232.
- Deno, S.L. (1986). Formative evaluation of individual student programs: A new role for school psychologists. *School Psychology Review, 15*, 358-374.
- Ehly, S., Reimers, T.M., & Keith, T.Z. (1986). Discriminant validity of the Personality Inventory for Children: Can it identify learning disabled children? *Learning Disabilities Research, 2*, 26-31.
- Elliott, S.N., Sheridan, S.M., & Gresham, F.M. (1989). Assessing and treating social skills deficits: A case study for the scientist-practitioner. *Journal of School Psychology, 27*, 197-222.
- Gerber, M.M. (1988). Tolerance and technology of instruction: Implications for special education reform. *Exceptional Children, 54*, 309-314.
- Good, R.H., & Salvia, J. (1988). Curriculum bias in published, norm referenced reading tests: Demonstrable effects. *School Psychology Review, 17*, 51-60.
- Gresham, F.M. (1988). Review of the Revised Children's Manifest Anxiety Scale. In J.C. Conoley & J.J. Kramer (Eds.), *The Tenth Mental Measurements Yearbook* (pp. 625-630). Lincoln, NE: University of Nebraska Press.

- Gresham, F.M., & Elliott, S.N. (1984). Assessment and classification of children's social skills: A review of methods and issues. *School Psychology Review, 13*, 292-301.
- Gresham, F.M., & Elliott, S.N. (1989). *Social Skills Rating System*. Circle Pines, MN: American Guidance Service.
- Gresham, F.M., & Reschly, D.J. (1988). Issues in the conceptualization, classification and assessment of social skills in the mildly handicapped. In T.R. Kratochwill (Ed.), *Advances in School Psychology*, Vol. 6 (pp. 203-247). Hillsdale, NJ: Lawrence Erlbaum.
- Guilford, J.P. (1967). *The nature of human intelligence*. New York: McGraw-Hill.
- Harrison, P.L. (1989). Adaptive behavior: Research to practice. *Journal of School Psychology, 27*, 301-317.
- Harrison, P.L., & Kamphaus, R.W. (1984, April). *Comparison between the K-ABC and Vineland Adaptive Behavior Scales*. Paper presented at the meeting of the National Association of School Psychologists, Philadelphia.
- Heath, C.P., & Obrzut, J.E. (1986). Comparison of three measures of adaptive behavior. *American Journal of Mental Deficiency, 89*, 205-208.
- Hilliard, A.G. (1979). Standardization and cultural bias as impediments to the scientific study and validation of "intelligence." *Journal of Research and Development in Education, 58*, 47-58.
- Horn, J.L. (1985). Remodeling old models of intelligence. In B.B. Wolman (Ed.), *Handbook of Intelligence* (pp. 267-300). New York: Wiley.
- Horn, J.L. (1986). Some thoughts about intelligence. In R.J. Sternberg & D.K. Detterman (Eds.), *What is intelligence? Contemporary viewpoints on its nature and definition* (pp. 91-96). Norwood, NJ: Ablex.
- Humphreys, L.G. (1985). Race differences and the Spearman hypothesis. *Intelligence, 9*, 275-283.
- Jackson, G.D. (1975). Another psychological view from the Association of Black Psychologists. *American Psychologist, 30*, 88-93.
- Jenkins, J.R., Deno, S.L., Mirkin, P.K. (1979). Measuring pupil progress toward the least restrictive environment. *Learning Disability Quarterly, 2*, 81-92.
- Jenkins, J.R., & Pany, D. (1978). Curriculum biases in reading achievement tests. *Journal of Reading Behavior, 10*, 345-357.
- Jensen, A.R. (1984). The black-white difference on the K-ABC: Implications for future tests. *Journal of Special Education, 18*, 377-408.
- Jensen, A.R. (1985). The nature of the black-white difference on various psychometric tests: Spearman's hypothesis. *Behavioral and Brain Sciences, 8*, 193-263.
- Jensen, A.R. (1980). *Bias in mental testing*. New York: Free Press.
- Kaplan, R.M. (1985). The controversy related to the use of psychological tests. In B. Wolman (Ed.), *Handbook of Intelligence* (pp. 465-504). New York: Wiley.
- Kaufman, A.S. (1976). A new approach to the interpretation of test scatter on the WISC-R. *Journal of Learning Disabilities, 9*, 160-168.
- Kaufman, A.S. (1979). *Intelligent testing with the WISC-R*. New York: Wiley.
- Kaufman, A.S., Kamphaus, R.W., & Kaufman, N.L. (1985). New directions in intelligence testing: The Kaufman Assessment Battery for Children (K-ABC). In B. Wolman (Ed.), *Handbook of Intelligence* (pp. 663-698). New York: Wiley.

- Kaufman, A.S., & Kaufman, N.L. (1977). *Clinical evaluation of young children with the McCarthy Scales*. New York: Grune & Straton.
- Kaufman, A. S., & Kaufman, N.L.(1983). *K-ABC: Kaufman Assessment Battery for Children: Interpretive manual*. Circle Plaines, MN: American Guidance Service.
- Keith, T.Z. (1985). Questioning the K-ABC: What does it measure? *School Psychology Review*, 14, 9-20.
- Keith, T.Z. (1986). Factor structure of the K-ABC for referred school children. *Psychology in the Schools*, 23, 241-246.
- Keith, T.Z. (1987). Assessment research: An assessment and recommended interventions. *School Psychology Review*, 16, 276-289.
- Keith, T.Z., & Dunbar, S.B. (1984). Hierarchical factor analysis of the K-ABC: Testing alternate models. *The Journal of Special Education*, 18, 367-375.
- Keith, T.A., Fehrmann, P.G., Harrison, P.L., & Pottebaum, S.M. (1987). The relationship between adaptive behavior and intelligence: Testing alternative explanations. *Journal of School Psychology*, 25, 31-43.
- Kelly, E.J. (1988). Personality Inventory for Children: Selected scales in differentiating conduct-disordered and emotionally disturbed children. *Psychological Reports*, 63, 395-401.
- Keogh, B.K. (1988). Improving services for problem learners: Rethinking and restructuring. *Journal of Learning Disabilities*, 21, 19-22.
- Knoff, H.M. (1989). Review of the personality inventory for children. In J.C. Conoley & J.J. Kramer (Eds.), *The Tenth Mental Measurements yearbook* (pp. 625-630). Lincoln, NE: University of Nebraska Press.
- Lachar, D. (1982). *Personality Inventory for Children (PIC): Revised format manual supplement*. Los Angeles: Western Psychological Services.
- Larry P. et al. v. Wilson Riles et al., C712270 (United States District Court for the Northern District of California, October 1979, slip opinion).
- Leinhardt, G., & Seewald, A.M. (1981). Overlap: What's tested, what's taught? *Journal of Educational Measurement*, 18, 85-96.
- Lieberman, L.M. (1985). Special education and regular education: A merge made in heaven. *Exceptional Children*, 51, 513-516.
- Marston, D., Lowry, L., Deno, S.L., & Mirkin, P.K. (1981). *An analysis of learning trends in simple measures of reading, spelling, and written expression: A longitudinal study* (Research Report No. 49). Minneapolis: University of Minnesota Institute for Research on Learning Disabilities.
- Martin, R.P. (1988). *Temperament Assessment Battery for Children: Manual*. Brandon, VT: Clinical Psychology.
- Mattison, R.E., Bagnato, S.J., & Brubaker, B.H. (1988). Diagnostic utility of the Revised Children's Manifest Anxiety Scale in children with DSM-III anxiety disorders. *Journal of Anxiety Disorders*, 2, 147-155.
- Mayr, E. (1982). *The growth of biological thought*. Cambridge, MA: Bknap Press.
- Mealor, D.J., & Richmond, B.O. (1980). Adaptive behavior: Teachers and parents disagree. *Exceptional Children*, 46, 386-389.
- Mehrens, W.A. (1984). National tests and local curriculum: Match or mismatch? *Educational Measurement: Issues and Practices*, 3, 9-15.
- Mehrens, W.A., & Phillips, S.E. (1986). Detecting impacts of curricular differences in achievement test data. *Journal of Educational Measurement*, 23, 185-196.

- Mercer, J.R., & Lewis, J.F. (1977). *System of Multicultural Pluralistic Assessment: Parent interview manual*. New York: Psychological Corp.
- Naglieri, J.A. (1985). *Matrix Analogies Test: Expanded Form*. Columbus, OH: Charles E. Merrill.
- Naglieri, J.A., & Jensen, A.R. (1987). Comparison of black-white differences on the WISC-R and the K-ABC: Spearman's hypothesis. *Intelligence, 11*, 21-43.
- Obrzut, J.E., & Cummings, J.A. (1983). The projective approach to personality assessment: An analysis of thematic picture techniques. *School Psychology Review, 12*, 414-420.
- PASE: Parents in action on special education et al. v. Hannon et al. No. 74 C3586 (United States District Court for the Northern District of Illinois, Eastern Division, July 1980, slip opinion).
- Reynolds, C.R. (1982). The problem of bias in psychological assessment. In C.R. Reynolds, & T.B. Gutkin (Eds.), *The handbook of school psychology* (pp. 178-208). New York: Plenum Press.
- Reynolds, C.R., & Kaufman, A.S. (1985). Clinical assessment of children's intelligence with the Wechsler Scales. In B. Wolman (Ed.), *Handbook of Intelligence* (pp. 601-661). New York: Wiley.
- Reynolds, C.R., & Paget, K.D. (1983). National normative and reliability data for the Revised Children's Manifest Anxiety Scale. *School Psychology Review, 12*, 324-336.
- Reynolds, C.R., & Richmond, B.O. (1976). What I think and feel: A revised measure of children's manifest anxiety. *Journal of Abnormal Child Psychology, 6*, 271-280.
- Reynolds, C.R., & Richmond, B.O. (1985). *Revised Children's Manifest Anxiety Scale*. Los Angeles: Western Psychological Services.
- Reynolds, M.C., Wang, C., & Walberg, H.J. (1987). The necessary restructuring of special and regular education. *Exceptional Children, 53*, 391-398.
- Saigh, P.A. (1989). The validity of the DSM-III posttraumatic stress disorder classification as applied to children. *Journal of Abnormal Psychology, 98*, 189-192.
- Salvia, J., & Ysseldyke, J.E. (1988). *Assessment in special and remedial education* (4th ed.). Dallas: Houghton Mifflin.
- Sattler, J.M. (1988). *Assessment of children* (3th ed.). San Diego, CA: Jerome M. Sattler.
- Schmidt, W.H. (1983). Content biases in achievement tests. *Journal of Educational Measurement, 20*, 165-178.
- Shinn, M.R., & Marston, D. (1985). Differentiating mildly handicapped, low-achieving and regular education students: A curriculum-based approach. *Remedial and Special Education, 6*, 31-45.
- Sparrow, S.S., Balla, D.A., & Cicchetti, D.V. (1984). *Manual for the Vineland Adaptive Behavior Scales, Interview Edition, Survey Form*. Circle Pines, MN: American Guidance Service.
- Spearman, C. (1927). *The abilities of man*. New York: MacMillan.
- Stainback, S., & Stainback, W. (1984). A rationale for the merger of special and regular education. *Exceptional Children, 51*, 102-111.
- Sternberg, R.J. (1984). The Kaufman Assessment Battery for Children: An information-processing analysis and critique. *Journal of Special Education, 18*, 269-279.

- Sternberg, R.J. (1985a). *Beyond IQ*. New York: Cambridge University Press.
- Sternberg, R.J. (1985b). Cognitive approaches to intelligence. In B. Wolman (Ed.), *Handbook of Intelligence* (pp. 59-118). New York: John Wiley.
- Sternberg, R.J. (1986, Fall). The future of intelligence testing. *Educational Measurement: Issues and Practices*, 10-22.
- Sternberg, R.J., & Salter, W. (1982). Conceptions of intelligence. In R.J. Sternberg (Ed.), *Handbook of human intelligence* (pp. 3-28). New York: Cambridge University Press.
- Thorndike, R.L., Hagen, E.P., & Sattler, J.M. (1986). *Technical manual: Stanford Binet Intelligence Scale*, 4th Ed.. Chicago: Riverside.
- Wall, S.M., & Paradise, L.V. (1981). A comparison of parent and teacher reports of selected adaptive behaviors of children. *Journal of School Psychology*, 19, 73-77.
- Wechsler, D. (1944). *The measurement of adult intelligence*. Baltimore: Williams & Wilkins.
- Weinberg, R.A. (1989). Intelligence and IQ: Landmark issues and great debates. *American Psychologist*, 44, 98-104.
- Weiss, D.J. (1982). Improving measurement quality and efficiency with adaptive testing. *Applied Psychological Measurement*, 6, 473-492.
- Weiss, D.J., (Ed.). (1983). *New horizons in testing: Latent trait test theory and computerized adaptive testing*. New York: Academic Press.
- Weiss, D.J., & Kingsbury, G.G. (1984). Application of computerized adaptive testing to educational problems. *Journal of Educational Measurement*, 21, 361-375.
- Will, M.C. (1986). Educating children with learning problems: A shared responsibility. *Exceptional Children*, 52, 411-415.
- Williams, R.L. (1974). From dehumanization to black intellectual genocide: A rejoinder. In G.J. Williams & S. Gordon (Eds.), *Clinical child psychology: Current practices and future perspectives*. New York: Behavioral Publications.
- Wirt, R.D., Lachar, D., Klinedinst, J.K., & Seat, P.D. (1984). *Multidimensional description of child personality: A manual for the personality inventory for children* (Revised 1984 by D. Lachar). Los Angeles: Western Psychological Services.
- Witt, J., & Elliott, S. (1986). Acceptability of classroom interventions. In T.R. Kratochwill (Ed.). *Advances in school psychology*, Vol. 4. Hillsdale, NJ: Lawrence Erlbaum.
- Woodcock, R.W. (1989, August). *Factor structure of the tests of cognitive ability from the 1977 and 1989 Woodcock-Johnson*. Paper presented at the ACER seminar on Intelligence. Melbourne, Australia.
- Ysseldyke, J.E., & Christenson, S.L. (1987). *The Instructional Environment Scale: A comprehensive methodology for assessing an individual student's instruction*. Austin, TX: Pro-Ed.

Jack A. Cummings, Associate Professor and Director of the School of Psychology Program at Indiana University (Bloomington), has conducted research on variables associated with psychologists' decisions to place children in special education classes, as well as criterion validity investigations of cognitive ability measures. The development and evaluation of interactive video programs for preservice psychologists to refine their interviewing skills have been the focus of his efforts for the past two years.

Marc C. Laquerre is a doctoral student in school psychology at Indiana University (Bloomington). His former studies were at College of the Holy Cross, Worcester, Mass. Mr. Laquerre's research interests include visual-motor assessment, adolescent substance abuse, and therapeutic interventions with families and children.

Jack A. Cummings, professeur agrégé et directeur du programme de psychologie scolaire à l'université d'Indiana (Bloomington), a effectué des recherches sur les variables qui entrent dans les décisions prises par les psychologues de placer certains enfants dans des programmes d'enseignement spécial et sur la validité des critères de mesure des facultés cognitives. Depuis deux ans, il travaille à la conception et à l'évaluation de programmes vidéo interactifs en vue de permettre aux psychologues de peaufiner leurs techniques d'entrevue avant l'emploi.

Marc C. Laquerre est étudiant de doctorat en psychologie scolaire à l'Université d'Indiana (Bloomington). Il a fait ses études au College of the Holy Cross, à Worcester (Mass.). Ses recherches portent sur l'évaluation visuelle motrice, l'abus de substance par les adolescents et les interventions thérapeutiques auprès des familles et des enfants.

