DOCUMENT RESUME

ED 279 720 TM 870 166

AUTHOR Flaitz, Jim
TITLE Why Intelligence Tests Must Change.
PUB DATE 20 Nov 86
NOTE 14p.; Paper presented at the Annual Meeting of the Mid-South Educational Research Association (Memphis, TN, November 19-21, 1986).
PUB TYPE Speeches/Conference Papers (150) -- Viewpoints (120) -- Reports - Descriptive (141)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Brain Hemisphere Functions; Cognitive Ability; *Cognitive Measurement; Diagnostic Tests; *Intelligence Tests; Measurement Techniques; Piagetian Theory; *Testing Problems

ABSTRACT

Today's tests of intelligence are largely unchanged over the past 70 to 80 years, despite substantial changes in the way intelligence is conceptualized. The history of intelligence testing reveals that much more has been done to perfect the measurement of traits that are static and immutable than has been done to make or keep intelligence tests relevant. The major factors that make today's intelligence tests suspect are the advances in theory, especially so in the area of cognitive psychology; more extensive needs of clinicans; and changes in the populations with which intelligence tests are being used. Tomorrow's tests will have to be more theory-based, produce better diagnostic information, and be relevant to the types of test-takers with which they will increasingly be used. However, changes in tests will not occur overnight, as publishers will not move rapidly in the absence of a proven market, and the market will not arise until test users are widely educated to the deficiencies of the existing tests and the promise of tomorrow's tests. (Author)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED279720

WHY INTELLIGENCE TESTS MUST CHANGE

Jim Flaitz
Univerisity of Southwestern Louisiana

A paper presented at the annual meeting of
the Mid-South Educational Research Association
November 20, 1986

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
J. Flaitz

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

WHY INTELLIGENCE TESTS MUST CHANGE

Over the past 20 years approximately, a significant number of developments have occurred in the fields of human psychology and psychometry which, taken collectively suggest that we may be due for some major changes in the way we conceptualize, and more importantly the way we measure intelligence. It is remarkable that the underlying premises upon which the measurement of intelligence has been based continue to hold sway, given the dramatic increases in our knowledge of cognitive processes. In fairness, the traditional measures of intelligence have served long and well, and will doubtless continue to be used and defended for some time to come. Nevertheless, the forces for change are there. The needs of clinicians have changed, as they are increasingly being required to provide specific and substantiated recommendations for intervention based upon their assessments. Contemporary thought and theory regarding intelligence is making traditional approaches to assessing intelligence increasingly less relevant. Where once psychometry drove theory, in the future theory will have to drive psychometry.

The remainder of this paper will examine the origins of the traditional measures of intelligence, the factors that are making the traditional measures less relevant, the obstacles that will continue to stand in the way of change, and a brief summary.

A brief history of intelligence testing

Sir Francis Galton is generally credited with devising the first measurement systems which were intended to index differences in human intelligence (Kaufman, Clark, & Flaitz, in press). It was Galton's belief, predicated largely upon the work of his cousin, Charles Darwin, that within the human species graduations of abilities would be found as a consequence of genetic variations. Galton's tests of sensory-motor skills were taken as evidence of intellectual potency based on two arguments. Extending the philosophies of Locke and Hume, Galton reasoned that all knowledge of the world derived from sensation. Hence the intelligence of an individual would manifest itself in an manner commensurate with the efficacy of the sensory apparatus. The second argument was based on the patterns observed in Galton's data. A close correlation could shown between his measures and independent measures of intellectual status such as the station of the individual in society. Galton's methods remained pre-eminent in England and Europe through the turn of the century, and they were brought to the U.S. by James McKeen Cattell.

In France, Alfred Binet succeeded in mounting a challenge to the type of measures Galton advocated, based upon his work for the French Minister of Public Instruction. Binet's goal was to implement a testing program to identify the academically deficient among the children entering the public school system. Binet devised a battery of assessment activities which tapped general fund of knowledge, verbal skills, memory, and reasoning (Binet & Simon, 1895). These measures were graduated in complexity and difficulty to be appropriate for various age groups. The instrument was intended to be interpreted as an index of academic readiness.

For a brief period at the turn of the century, the testing movement lost momentum as the result of a couple of widely reported studies. Wissler (1901) and Sharp (1898-99) investigated the measures of Binet and Galton and found little empirical evidence of a relationship between those measures and other indicators of academic orintellectual ability. In time, deficiencies in these studies associated with sample size and restriction of range were recognized, buth the general enthusiasm for testing was dampened until the outset of World War I. At that time, the need arose for efficient methods of selecting among newly inducted members of the armed forces. While Binet's scales were the more widely accepted measures of ability, the one-to-one method of administration was not well suited to the enormity of the task. Otis is credited with the development of a paper-and-pencil form of the Binet scales. Also during this process, the problem of illiterate and non-English speaking test-takers surfaced. Non-verbal scales (subsequently known as Army Beta) were devised to assess the general intelligence of recruits who could not be validly assessed with the paper-and-pencil test form (Army Alpha). Following the war, the use of intelligence tests, often virtually unchanged from the Army forms, proliferated in business and education (Kaufman & Flaitz, in press).

A significant milestone in the history of intelligence testing came about as a result of the work of David Wechsler. Wechsler, who was primarily a clinician by training, believed that while intelligence was a unitary trait, it was manifested in diverse forms, and important clinical information could be derived from a battery that assessed verbal and non-verbal aspects of intelligence separately. The result was the Wechsler Intelligence Scale for Children (WISC) and the Wechsler Adult Intelligence Scale (WAIS). Both scales have undergone revision, as has the Stanford-Binet, an American version of the Binet scales.

Shortcomings of traditional measures of intelligence

Both the Wechsler scales and the Stanford-Binet have their roots in the scales of Alfred Binet, dating back to the turn of the century. Binet evidenced some reluctance to describe his scales as measures as intelligence, although enthusiasts showed no similar reservation. Lewis Terman may be more responsible than anyone else for inextricably linking Binet's scales with the concept of intelligence.

Perhaps ironically, it may be that the available instrumentation did more to drive theorizing about intelligence than theory has ever done to shape the nature of the measures. The circumstances led Boring to offer the well known observation that "intelligence is what intelligence tests measure" (Boring, 1923).

The theoretical underpinnings of the traditional measures of intelligence, such as they are, existed at the time that Binet undertook the development of his first scales nearly 90 years ago. After Binet's scales had been in use for several years, Charles Spearman (1904) outlined a general theory of intelligence that proposed the existence of the g factor to account for the close correlation of apparently distinct measures of ability, such as Binet's scales on one hand and Galton's scales on the other. Spearman's theory actually served more to validate the measures of Galton and Binet than to influence them. The two factor theory of intelligence offered by Spearman thus can be seen as an attempt to account for the existing data produced by extant scales, rather than serving as the catalyst for instrument development.

Some years later, Thurstone (1938) suggested an alternative theory of intelligence in response to apparent statistical anomalies in Spearman's work. Thurstone chose to employ a factor analytic methodology in studying the measurement of intelligence, and he proposed that more than one factor might exist to explain the shared variance among a set of scales. Thurstone subsequently identified nine factors which he believed to represent primary abilities.

Raymond B. Cattell (1963) sought to resolve the apparent discrepancies existing between Spearman's two factor theory and Thurstone's primary abilities model. He felt this resolution could be demonstrated when second-order factor analysis was applied to the primary ability factors, with the emergence of a single super-ordinate factor akin to Spearman's g. Cattell went on to elaborate his theory which emerged as the theory of fluid and crystallized intelligence.

Thurstone was responsible for developing a number of measures including the Primary Abilities Test, and Cattell and Horn have produced a number of measures of ability, and yet none of these have successfully crossed over from the experimental lab, where they were developed, to the general marketplace.

Guilford (1979) has proposed a far more ambitious theory, the Structure-of-Intellect Model, which suggests that intelligence exists in three dimensions- operations, content, and products. With 5 identified types of operations, 4 types of content, and 6 types of products, Guilford's model ultimately posits 150 types of intelligence. Guilford and his associates have succeeded in developing instruments which assess some of the 150 categories, but the tests have generated little commercial interest.

While there are other important theoretical models of intelligence, the point is established that the traditional, commercially successful measures of intelligence are in no way based upon or closely related to any major theoretical position. Moreover, these intelligence tests have not changed significantly for approximately 80 years.

Of even greater concern than the observation that traditional tests of intelligence were not initially formulated along explicit theoretical lines is the recognition of how far research and theory have moved since these tests were first established. Perhaps most central to the new thrust of theory into the topic of intelligence is the treatment of intelligence as a flexible and fluid process or processes which develop over the lifespan of the individual. Also of importance has been the attempt to link cognitive processes with brain physiology. Among the theorist and researchers who have actively sought to introduce this alternative view of intelligence are Piaget, Horn and Cattell, and Das, Kirby, and Jarman.

Without elaborating on a major theoretical position, it can be suggested that Piaget's conceptualization of intelligence differed significantly from the traditional, static model in intelligence in several ways. Rather than treating intelligence as an innate capacity factor, Piaget (1972) viewed intelligence as a developmental process, grounded in an invariant sequence of stages during which intellectual skill was manifested in a quite distinctive manner, limited by inherent characteristics of the stage. What is implicit in Piaget's work is the suggestion that intelligence is not the same things at all stages of an individual's life, and that while the processes that ultimately dictate the progressive unfolding of intellectual functioning may be deterministic, the traditional methods of measuring intelligence cannot adequately deal with the dramatically different manifestations of intelligence which arise at various stages of intellectual development. More recently, numerous researchers (e.g., Commoner, et al., 1982; Kitchener, 1983; Kramer, 1983) have offered the opinion that for adults it is necessary to posit a stage of cognitive functioning beyond Piaget's formal operations; a stage characterized by a new form of flexibility in dealing with problems having no simple and scientific solution. Again the implication is that if the intellectual functioning of the individual is to be appropriately gauged, the instrumentation of that assessment must in some way incorporate mechanisms for addressing the distinctive nature of intelligence at any given stage.

Horn and Cattell (1966), whose theory of fluid and crystallized intelligence has already been alluded to, represent another important departure from traditional approaches to assess intellectual functioning. In an article dealing with the issue of culture fair testing, Cattell (1979) points out that traditional tests of intelligence have been and continue to be the assessment tools of choice based upon their effectiveness in predicting academic achievement, but that this application is ultimately a very narrow one. Following World Wars I and II, traditional tests of intelligence did a poor job of predicting how well an adult leaving school would do in learning to fire a naval gun or fly a plane. Cattell ascribed this failing to the absence of any fluid intelligence component to the traditional intelligence tests, and suggested that the 20% of purely scholastic knowledge variance found in traditional tests of intelligence would actually reduce predictive validity in such situations, as it introduced misleading, systematic error.

Das, Kirby, and Jarman (1975) have advanced a comprehensive theory of intelligence and cognitive functioning based upon a model of information processing which can be broadly linked to specific cortical areas. Their theory is rooted in the work of Luria (1970), who observed that the cortex is engaged in two types of integrative activity- simultaneous and successive. Luria divided the brain into three maor functional units- the arousal and attention unit located in the upper brain stem and reticular formation; the input, recoding, and storage unit located in the occipital, parietal, and frontal-temporal regions; and the unit for planning and programming behavior, located in the frontal lobe. Extensive research with the model lead Das, Kirby, and Jarman to conclude that a legitimate simultaneous-sequential dichotomy in processing exists, and that the more traditional models that treat rote memory (which depends heavily upon sequential processing) as subordinate to problem solving (which implies simultaneous processing) are incorrect. For assessment, the implication of the work of Das, Kirby, and Jarman (and by extension, Luria) is that to appropriately measure intelligence, an instrument must possess the capacity to assess simultaneous and sequential processes separately. Their work has suggested that these two processes adequately account for variations in cognitive functioning in children; however, in adults, an assessment instrument would also need a third component to address the planning process, which appears to manifest itself (in Piagetian-like fashion) only as the individual matures.

In the area of neuropsychology, the work of Bogen (1969), Sperry (1968), Luria (1970), and Gazzaniga (1970) has proven invaluable in understanding the manner in which the brain functions and is integrated in those functions. Beginning with Sperry's split brain studies, it has become evident that in some fashion the brain functions in two distinct modalities. While it has been suggested that this dichotomy represents a left brain-right brain functioning (Bogen, 1969; Sperry, 1968) Luria has suggested that the divisions of the brain are more complex. Whatever the ultimate resolution of the problem of location of specific brain function, it is clear that the methodology exists for meaningfully studying the linkage between brain structure and brain function, and that assessment tools for the measurement of brain function must in some fashion acknowledge the growing body of knowledge concerning the function-structure relationship.

When the Binet scales were first developed, their purpose was simply to identify children who were slow or not prepared for school. As such, the major evidence for their validity stemmed from the ability of the scales to predict academic achievement. At that time there was little concern that the instrument might itself be heavily loaded on academic achievement. Wechsler produced his scales to serve a slightly different purpose, namely to assist in the clinical diagnosis of intellectual impairment (Wechsler, 1975). To that extent, the Wechsler scales might be said to come closer to the theoretical ideal of intelligence than do the Binet scales. However, even in the case of the Wechsler scales, their use has most often been as a means of confirming or disconfirming an hypothesis of dysfunction, or as a mechanism for the identification of the source of dysfunction.

Events of the past decade have presented new challenges to the clinician attempting to employ intelligence tests. There are those who call for the abolition of all intelligence tests, citing problems of abuse and bias. On the other hand, many clinicians are obliged to utilize these same tests in rigid and mechanical fashion, ignoring the known limits of the tests in order to comply with bureaucratic mandates. In part the problem is one of practice; that is, the clinician must have superior powers of inference and deduction than the instruments he uses. But the clinician is nonetheless limited to the extent that the tests he has available provide little information beyond the ability to distinguish function from dysfunction, and to produce a label for the dysfunction. The findings of clinical assessment must lead to some specific intervention which will benefit the client, and must itself generate information that suggests appropriate intervention strategies. In large part this requirement of clincal assessment has not been met by traditional tests of intelligence.

One last source of difficulty for the traditional measures of intelligence relates to their apppropriateness for use with adults. The Binet scales were created for the express purpose of assessing functioning of children. The extension of the test to young adult and/or adult populations represents little more than an act of faith predicated on the tenet that intelligence is static, or at most accretional. While the Wechsler scales were developed in such as manner as to result in separate instruments for children and adults, examination of the subtests of each form reveal a remarkable degree of overlap. Again, this is presumably due to the traditional view of intelligence as relatively immutable and unchanging over the lifespan. The major concessions to the delineation of child and adult intelligence on the two forms of the Wechsler scales appear to be a slight redefinition of some of the tasks (typically making them more difficult) and the provision of age-appropriate norms. As was earlier noted, Cattell addressed the problem of assessing adult intelligence, pointing out the disappointing performance ofptraditional scales in predicting outcomes that are more relevant for adults than academic achievement.

The question of developmental changes in intellectual potency in the later adult years has created controvery as well. The importance of the question grows as the U.S. population becomes older. The traditional view, based on trends obtained from cross-sectional analysis of intelligence test scores holds that following late adolescence the intellectual function as represented by a graph line first plateaus and then declines in the later years. Interpretation of this data has differed though. One interpretation is that intellectual potency declines with age, while an alternative hypothesis suggests that cohort differences in cross-sectional studies, and intervening history in longitudinal studies are responsible for the appearance of these differences (Schaie, 1972). However, both camps seem to agree that the instruments typically employed in the study of intellectual functioning are poorly suited to the measurement of intelligence in older adults.

The future of intelligence testing

The sorts of changes needed in the intelligence tests of the future are readily inferred from the limitations of today's tests. They must be more theory-based, and specifically must address theoretical issues that have arisen over the past twenty years. Given that much of the contemporary work in the area of intelligence has tended to conceptualize intelligence as process-based, it follows that measures of intelligence in the future must focus on the processes of cognitive functioning rather than the supposed product of those processes. Another feature of tomorrow's intelligence tests will hopefully be a more relevant contextual basis, especially so for measures intended to assess adult intelligence. That is to say, the nature of the items and tasks of the test should be related to the types of tasks and circumstances more typical of adult life than is true at present. Because tomorrow's tests must be more clinically useful, it is likely that they will somehow generate, more readily, diagnostic information that can be applied to intervention. Thus, the intelligence test used in the academic setting would ideally tell the instructor important information regarding the test-takers potential for learning specific types of skills. Finally, it may be that more tests, designed for relatively narrowly defined purposes, will emerge and achieve co-equal prominence with tests intended to produce global assessments of intelligence.

The most complete example of tomorrow's tests available today may be the Kaufman Assessment Battery for Children (K-ABC). The K-ABC is theory-based, utilizing the principles of several theories (e.g., Das, Kirby, and Jarman, 1975; Luria, 1970; and Horn and Cattell, 1966). It is process-oriented. And perhaps of greatest significance, the instrument yields data that can guide the teacher to productive activities which will benefit the child. Other instruments which share the features of theory-based, process-oriented, and diagnostically useful are under development (e.g., Sternberg, 1985; Feuerstein, 1979; Gardner, 1983; Hunt,1978). Additionally, significant strides are already being made in the computerization of certain aspects of intelligence testing, which may result in more efficient testing.

Obstacles to changes in the traditional tests of intelligence

While there is much reason for optimism in the area of intelligence testing in the future, certain factors will impede the rapid restructuring of the field. The foremost obstacle would appear to be the test publisher. The history of intelligence testing seems to reveal nearly as much about test publishing as it does about test development. Test publishers understandably have a major investment in the tests that are currently popular, and will be hard pressed to either abandon those tests or to even-handedly promote newer and potentially more promising tests. Factors of cost and market control similarly make it difficult for new publishers to appear on the scene offering superior products for competitive prices. Thus it is clear that publishers will be a force of inertia for the foreseeable future, just as they have been in the past.

A related obstacle arises from the habits of today's test users. Publishers respond to market demand, and if a remarkable new demand for new tests were to emerge, more than one publisher would undoubtedly respond to the demand. However, the current users of tests are to a large extent content with the tests currently available to them. Perhaps more accurately, they have learned to accommodate themselves to the existing shortcomings in today's tests (or are unaware of those shortcomings). Whatever the case until test users want better tests, the publishers are unlikely to begin supplying new tests. The key to what tomorrow's test users will demand may lie with the programs that train those individuals. Like Thomas Kuhn's description of change in science, it may be that the change in habits of test users will be based more upon attrition than upon a broad-based change in attitude among existing users.

Summary

Today's tests of intelligence are largely unchanged over the past 70 to 80 years, despite substantial changes in the way intelligence is conceptualized. The history of intelligence testing reveals that much more has been done to perfect the measurement of traits that are static and immutable than has been done to make or keep intelligence tests relevant.

The major factors that make today's intelligence tests suspect are the advances in theory, especially so in the area of cognitive psychology, more extensive needs of clinicians, and changes in the populations with which intelligence tests are being used. Tomorrow's tests will have to be more theory-based, produce better diagnostic information, and be relevant to the types of test-takers with which they will increasingly be used. However, changes in tests will not occur overnight, as publishers will not move rapidly in the absence of a proven market, and the market will not arise until test users are widely educated to the deficiencies of the existing tests and the promise of tomorrow's tests.

References

Bogen, J.E. (1969). The other side of the brain: Parts I, II, & III. *Bulletin of the Los Angeles Neurological Society*, *34*, 73-105; 135-162; 191-203.

Boring, E.G. (1923, June 6). Intelligence as the tests test it. *The New Republic*, 35-37.

Cattell, R.B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, *54*, 1-22.

Cattell, R.B. (1972). Are culture fair intelligence tests possible and necessary? *Journal of Research and Development in Education*, *12* (2), 3-13.

Commons, M.L., Richards, F.A., & Kuhn, D. (1982). Case for levels of reasoning beyond Piaget's stage of formal operations. *Child Development*, *53*, 1058-1069.

Das, J.P., Kirby, J., & Jarman, R.F. (1975). Simultaneous and successive syntheses: An alternative model for cognitive abilities. *Psychological Bulletin*, *82*, 87-103.

Feuerstein, R. (1979). *The dynamic assessment of retarded performers: The learning potential assessment device, theory, instruments and techniques*. Baltimore, MD: University Park Press.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences.* New York: Basic Books.

Gazzaniga, M.S. (1970). *The bisected brain*. New York: Appleton-Century-Crofts.

Guilford, J.P. (1979). Intelligence isn't what it used to be: What to do about it. *Journal of Research and Development in Education*, *12* (2), 33-46.

Horn, J.L. & Cattell, R.B. (1966). Refinement and test of the theory of fluid and crystallized ability intelligences. *Journal of Educational Psychology*, *57*, 253-270.

Hunt, E.B. (1978). Mechanics of verbal ability. *Psychological Review*, *85*, 109-130.

Kaufman, A.S., Clark, J. & Flaitz, J.R. (in press). Intelligence testing. In Bolton, B. (Ed.) *Handbook of measurement and evaluation in rehabilitation*. pp. 59-74. Baltimore, MD: Paul H. Brookes Publishing.

Kaufman, A.S. & Flaitz, J.R. (in press). Intellectual growth. In Hersen, M. & Van Hasselt, V.B. (Eds.) *Handbook of Adolescent Psychology*. New York: Pergamon Press.

Kitchener, K.S. (1983). Cognition, metacognition, and epistemic cognition. *Human Development*, *26*, 222-232.

Kramer, D.A. (1983). Post-formal operations? A need for further conceptualization. Human Development, 26, 91-105.

Luria, A.R. (1970). The functional organization of the brain. Scientific American, 222, 66-78.

Piaget, J. (1972). Intellectual evolution from adolescence to adulthood. Human Development, 15, 1-12.

Schaie, K.W. (1972). Limitations on the generalizability of growth curves of intelligence: A reanalysis of some data from the Harvard growth study. Human Development, 15, 141-152.

Sharp, S.E. (1898-99). Individual psychology: A study in psychological method. American Journal of Psychology, 10, 329-391.

Spearman, C. (1904). General intelligence, objectively determined and measured. American Journal of Psychology, 15, 201-293.

Sperry, R.W. (1968). Hemisphere deconnection and unity in conscious awareness. American Psychologist, 23, 723-733.

Sternberg, R.J. (1985). Beyond IQ: A triarchic theory of human intelligence. New York: Cambridge University Press.

Thurstone, L.L. (1938). Primary mental abilities. Chicago: The University of Chicago Press.

Wechsler, D. (1975). Intelligence defined and undefined. American Psychologist, 30 (2), 135-139.

Wissler, C. (1901). The correlation of mental and physical tests. Psychological Review, Monograph Supplement, 3, No.6.