

# ASSESSMENT OF INDIVIDUAL DIFFERENCES<sup>1,2</sup>

BY LEE J. CRONBACH

*Bureau of Educational Research, University of Illinois, Urbana, Illinois*

Assessment may be defined as the obtaining and evaluating of information regarding individual differences. Topics usually covered under the heading of "Individual Differences" are treated here. Space restrictions limit the review to major trends in the (estimated) 20,000 pages of relevant material considered.<sup>3</sup>

## THE FAILURE OF GLOBAL ASSESSMENT

Assessment in the OSS style has now been proved a failure. Earlier validation studies are reviewed by Symonds (164) and Kelly (99). In the Menninger project on assessment of 457 psychiatric trainees, neither tests nor psychiatric interviews had satisfactory validities (117). Holtzman & Sells (87) supplied projective data on student pilots to 19 prominent clinicians. The clinicians agreed with each other in identifying men eliminated on emotional grounds but could not agree with the criterion. In the Swedish Army, interviewers added negligibly to validity from tests [Husén (90)].

More favorable evidence is reported by Biesheuvel *et al.* (14). Among South African air cadets, psychomotor tests predicted pass-fail with validity .46. Ratings based on observation correlated .45 and the two yielded a multiple *R* of .62. An assessment of potential submariners correlated .28 with a later shipboard criterion in a group of restricted range [Mackie *et al.* (118)]. In both studies, ratings may have been based in part on ability rather than temperament.

Assessment encounters trouble because it involves hazardous inferences. Very little inference is involved when a test is a sample of the criterion or when an empirical key is developed. Simple test interpretations involve inference from test to construct to behavioral prediction. But assessors attempt a maximum inference from tests. As current writers describe the process (e.g., 9, 44, 100, 121, 124, 152), personality theory is applied to weave nomothetic constructs into a construct of the individual's personality structure; predictions are then derived by inferring how that structure will interact with the known or guessed properties of the situation. Assessors have been foolhardy to venture predictions of behavior in unanalyzed

<sup>1</sup> The survey of the literature pertaining to this review was completed in May, 1955.

<sup>2</sup> The following abbreviations are used in this chapter: MMPI (Minnesota Multiphasic Personality Inventory); PMA (Primary Mental Abilities); TAT (Thematic Apperception Test).

<sup>3</sup> Bibliographic assistance from Jack C. Merwin and Leonardo Ancona is gratefully acknowledged.

situations, using tests whose construct interpretations are dubious and personality theory which has more gaps than solid matter. Symonds (164) has this afterthought, "It is doubtful whether the TAT can be expected on theoretical grounds to predict any tangible overt expression of personality such as might be exemplified in academic success or work proficiency." And Luborsky (117) says, "Reviewing some predictions on which we erred, we were impressed with our correct assessment of many specific qualities and our inability to cast these up into proper balance so as to judge ability to develop skill as a psychiatrist."

The negative findings on maximum inference have to some extent divided psychologists between a conservative and an imaginative camp. Name-calling appears, as in Bellak's reference, while defending TAT, to "methodological bluenoses (who often mistake obsessive doubt for a scientific attitude and their intellectual sterility for caution and eclecticism)" (9, p. x). With poetic justice, the Fishers (56) employ a picture interpretation test to show that objectively-oriented psychology students are less insecure than those who are more intuitive. Simoneit (159) argues vehemently that objectification loses psychological meaning and that characterological study of the expressive behavior of the subject is required. Meili (125) counterattacks, arguing that *Ausdruckpsychologie* suffers from the same limitations as formal tests and has equal need for controlled research.

Others who recognize the schism speak more moderately. Cronbach (38) prevailed upon an anthropologist to describe dispassionately the national character of Clinicians and Psychometrikans. Rosenwald (145) criticizes research which seems "almost designed to devalue" clinical tests, but also shows that clinical reports are often made useless by unspecific descriptions and overliterary writing [see also McArthur (121)]. To dispel any belief that oceans separate measurer from *Versteher*, it remains only to cite this advice on assessment from that arch-measurer, Sir Cyril Burt: "The final case-summary should comprise, not merely analytic assessments in quantitative terms of the chief key-tendencies or 'factors,' but also a synthetic character-sketch, giving a qualitative picture of the individual, viewed as a dynamic, integrated, and unique human being" (23, p. 28). The critical question is why inferences go wrong, not whether inference is a legitimate mode of test interpretation.

#### THE CLINICAL PROCESS

*Value of clinical judgment.*—In experiments where clinicians and statistical formulas predict the same criterion, the formulas always do better [Meehl (124)]. Melton (126) found actuarial predictions of grade averages superior to those by counselors even when the counselors were given added data and the actuarial table itself. Halbower (75) obtained therapists' *Q*-sort descriptions of nine cases having the same MMPI pattern and used the modal sort as an actuarial key. This key was a better description for nine

new cases than was a sort made by a clinician directly from each individual's MMPI profile. This is the first use of actuarial methods to obtain a qualitative case description.

According to Meehl: An actuarial formula should be prepared for any recurrent decision. The unique function of the clinician is to predict micro-events, such that no statistical experience table can be prepared in time to be useful. The clinician can create new idiosyncratic constructs and thus predict the unprecedented. Such predictions are helpful during therapy, where no harm is done if many tries are erroneous because the subsequent interaction identifies the correct ones.

Even in the studies Meehl reviews, the clinical judge adds valid variance. But he makes overbold predictions, underweights relevant scores in favor of his clinical data, employs stereotypes instead of admitting his ignorance of what makes a good pilot or psychiatrist, and commits himself to a single "most likely" personality structure instead of reporting the tenable alternatives. These faults can be overcome. Wilson (178) interviewed subjects and estimated their scores on conventional tests of dominance, adjustment, and mental ability with validities .66 to .81. Cronbach (38, 39; see also 23) argues that certain assessment techniques have been unduly discredited. A procedure which covers many dimensions with low validity may have advantages over a single accurate test which answers only one question. Interviews and projectives have value for identifying which traits should be explored most thoroughly in the individual case.

*Testing as a social act.*—Test research has been dominated by the Galtonian view that the test is a sample of the subject's response to a standardized nonpersonal stimulus. Sarason (150) ably defends the contrary view that testing is a social interaction. Attitudes toward the tester, learned anxiety about tests, etc., modify the performance and the interpretation. Schafer's brilliant and absorbing chapter on "interpersonal dynamics in the test situation" (152) should be engraved on the heart of every tester. The following are examples of his provocative conclusions: such a test as the Wechsler should be administered in an unstandardized fashion; the role of the tester is "voyeuristic, autocratic, oracular, and saintly."

*The assessor's implicit personality theory.*—Bruner & Tagiuri (21) note that errors in interpretation of tests arise from the judge's implicit theory of trait interrelations. Walker (172) shows how such factors confuse communication in assessment reports; for example, psychologists regard sales effectiveness as dependent on verbal rather than nonverbal intelligence, but executives do not make this distinction. "Soundness" of graduate students is, among faculties, nearly synonymous with effectiveness, but psychologists emphasize warmth and de-emphasize eccentricity (5). Psychologists overweight introspectiveness in judging intelligence. Other evidence on this aspect of judgment comes from several studies (41, 69, 95, 96, 162).

## MODELS FOR ASSESSMENT

Eysenck discusses his views in several papers (50, 51, 53). At one pole he sees the Americans as trying to find ever more specific factors, a style which he attributes to their *S-R* tradition. At the other pole he places the Germans, whose typologies reflect a distrust of the quantitative. He believes that factor analytic dimensions may be organized hierarchically to describe types of persons. Burt (23) likewise accepts a hierarchy of traits. Eysenck emphasizes the factors of introversion and neuroticism. Among neurotics, the extrovert and introvert poles have their biological roots in Pavlov's inhibition and excitation, respectively. In supporting experiments by himself (49) and Franks (65), hysterics are less readily conditioned and more susceptible to figural after-effects than dysthymics (cf. 176).

Cattell (29) upholds simple structure against Eysenck's criterion rotation, but rejects the idea that factor analysis is presently ready for predictive use. Varied studies on individual differences [for example (82, 151, 181)] obtain data suited to Guttman's new radex model (73) for factor analysis, which may be preferable to arbitrary simple structure analysis.

Léon (105) comments on Russian reports to the Montreal Congress. The Russian view is that individual differences are developed by experience; appropriate remedial treatment of a motivated subject can radically alter specific abilities. The only biological verities are the Pavlovian dimensions (cf. Eysenck, above).

*Pattern and profile methods.*—Stephenson's efforts to use questionnaires to study each individual separately are well known. Technical reviews of his book (163) make clear that *Q*-methodology is treacherous. Cronbach & Gleser (40) praise the device of administering the same questions under varying instructions, but raise serious objections to the forced normal distribution, use of analysis of variance, and other features of the system. Eysenck (52) adds further criticisms.

Tiedemann & Bryan (167) demonstrate their centour technique, measuring how similar a student's interests are to typical successful students in each curriculum. The procedure involves the assumption that it is good for the student to enter the curriculum where he is typical. One may question the advantage of this approach over the fundamentally different multiple regression method.

*Idiosyncratic analyses.*—The Osgood "semantic differential" method of studying meaning systems is very promising for studying the individual case. Osgood & Luria (139) applied this method to the Thigpen-Cleckley case of triple personality, testing each personality in turn(!) and achieving a successful blind analysis.

Very similar ideas are elaborated by Kelly (100) in the first of two volumes on *The Psychology of Personal Constructs*. In his ingenious Role Construct Repertory Test, the subject first lists a large number of significant others in his life, and then lists attributes which distinguish certain individuals from others. This yields 15 scales important for this individual, on

which he describes all the significant others. The investigator studies the correlation among constructs, the perceived similarity among individuals, etc. The technique is indirect and exceptionally flexible but has serious technical deficiencies. The complexity of the data has lured Kelly's students into analyses so involved as to obscure serious errors in reasoning.

#### CRITERION-ORIENTED STUDIES: ABILITY

Anastasi (1) believes that traditional ability categories can be replaced by empirically defined group factors. Regression equations for differential tests against significant criteria, and interpretative rationale based on evidence are badly required, she adds. Terman (165) questions what specific ability factors add to prediction:

The experts in this field are divided into two schools. The British school, represented by Thomson, Vernon, and Burt, usually stop with the identification of at most three or four group factors in addition to *g*, while some representing the American school feed the scores of 40 or 50 kinds of tests into a hopper and manage to extract from them what they believe to be a dozen or fifteen separate factors. . . . It is to be hoped that further research will give us more information than we now have about the predictive value of the group factors. Until such information is available, the scores on group factors can contribute little to vocational guidance beyond what a good test of general intelligence will provide.

*Differential batteries.*—Berdie (11) studied four-year achievement in several college curricula. Vocational interest scores differentiated such groups, but differential abilities (PMA) did not predict field of concentration. Moreover, PMA was a less effective predictor of over-all grade average than were more general tests of aptitude or achievement.

A revised manual and technical supplement for PMA for ages 7 to 11 (166) shows little caution and much improper technical reporting. No evidence is reported for the dubious claim that the space ability is important in handcraft, art, geometry, etc., nor for the predictive advantage over mental age of the new "reading" and "arithmetic aptitude" scores. In contrast, new batteries by Holzinger & Crowder (88) and Segel & Raskin (157) are excellently presented. Striking variation is found in concurrent correlations with different achievement measures in the same field and with different samples. The tests often do not correlate where they are expected to correlate [see also Wolking (179) on PMA and the Differential Aptitude Tests]. In the Segel and Raskin data, arithmetic computation has the highest or nearly highest validity for virtually every high school subject. In shop, the highest validity is for paragraph meaning and one of the lowest for "applied science and mechanics." In neither battery does space show predictive value. We conclude that while factorial scores may be useful for a theory of abilities, as soon as testers make inferences to behavior in significant situations they encounter the same troubles as personality assessors. Group factors serve only when regression equations are constructed about the criterion in a single institution.

Flanagan's differential battery (59) measures 14 variables. Stanine scores for various occupations are based on formulas developed in an armchair. The printer is said to require superior judgment, comprehension, and expression; the same scores, supplemented by good memory(!), qualify one to be a professor of humanities. Of the two validation studies offered, the first is utterly irrelevant to this battery and the second is reported too sketchily to be meaningful. Publication with this manual is a step backward.

The doubtful practical status of differential testing is underlined by an announcement that the American Council on Education Test will be withdrawn from the market. This was the first fruit of multiple-factor research and sought to measure linguistic and quantitative "aptitudes." It will be replaced by the "Cooperative School and College Ability Test," which frankly measures verbal and quantitative achievement and is to be used to locate remedial needs as much as for prognosis.

*General intelligence tests.*—The important new general tests are the Lorge-Thordike Intelligence Tests for Kindergarten to Grade 12 (116), the individual Columbia Mental Maturity Scale (22), and the Wechsler Adult Intelligence Scale (174) which replaces the old Form I. The Adaptability Test (168) and the Short Employment Tests (10) are provided with good new manuals.

The nonverbal "semantic" test by Rulon & Schweiker (147) is validated against achievement of slow-learning Marines; it predicts appreciably better than the General Classification Test or Army Beta. Porteus (142) gives evidence for qualitative scoring of Maze performance. Wand & Mollenkopf (173) review 194 references on aptitude tests used for women.

### THE NATURE OF ABILITIES

*Cultural factors.*—Fahmy (54) tested children from the extremely restricted environment of the Southern Sudan and found ample evidence of cultural influences. The mean IQs on four tests were Goddard 74, Porteus 76, Goodenough 50, and Alexander 97. These children have no concept nor perception of the straight line and lack drawing experience. But the Alexander Passalong involves colored blocks, and colors are much emphasized in this culture. Biesheuvel (13) describes the testing of other primitive African groups, including performance testing for gold mine labor with pantomime motion picture instructions.

Twenty-eight American groups of differing social class were given tests by Haggard (74) under various conditions of motivation, practice, item form, etc. In general, the pupils of low status showed as much ability to learn as the "highs" and gained more under the conditions most familiar to them. Differences between lows and highs decreased under motivation and practice. Mere revision of items does not remove status differences; performance depends on the psychology of the testing situation.

*Coaching.*—In Britain, where tests at age 11 determine one's educational career, coaching has been thoroughly investigated. A symposium summa-

rized by Vernon (171) indicates that coaching plus practice on complete tests under standard conditions can raise IQs about 9 points. Practice alone has little effect. There is greater retest change among maladjusted children and on the more complex items. Holloway (84) studied the effect of the Thurstone Red Book on five-year-olds in preschool. The PMA IQ was raised significantly (13 points), but IQs on the Wechsler Scale for Children were unaffected. Hurtig (92) examined whether Progressive Matrices scores could be improved by supplementary explanation before retest. In the feeble-minded, changes up to 16 IQ points had been found. The main study is inconclusive because of poor statistical design but suggests that explanations are differentially helpful. According to Heim (80), the child will pass items, presented amid difficult items, which he fails in an easy context.

*Growth and development.*—Burt (24) examines whether abilities become more differentiated with age, severely criticizing the methods of past studies. In his own longitudinal data, his general factor accounts for much more variance at age 13 than at age 9. The statistical and semantic issues in this area are not, however, easily disentangled. Husén (91) found that twins have double the normal expectancy of falling in the lowest intelligence range. He suggests that this may be attributable to premature births among twins. The rest of his data, on 3000 cases, confirm previous research. Among mental defectives, Clarke & Clarke (35) found increases of 10 points in IQ after two years very common among subjects coming to the institution from bad homes.

*Factor studies.*—Zimmerman (181) hypothesizes that with increases in difficulty superficially similar spatial items will shift from measuring perceptual speed to spatial to visualization to reasoning. Using suitable reference tests, he confirms this except as regards reasoning.

The most significant contribution to the study of abilities since World War II is the program of Fleishmann & Hempel on psychomotor abilities. Information is available in numerous military reports and some journal articles, several of which fall in the current year (60 to 63, 81). In the interests of a comprehensive summary, we draw on an unpublished review of the program made available by Fleishmann. The work is intended to define the variables in motor tests and to determine the significance of these variables for prediction. Apparatus tests have considerable utility in predicting pilot success and school success of electricians, etc. Factorial studies covering dexterity (61), physical proficiency (81), simple psychomotor tests, and complex coordinations sharpen older factors and indicate new ones. Pencil-paper tests are not factorially similar to apparatus tests. These factor studies characterize almost every available psychomotor test in terms of general constructs, and new tests are devised to measure factors with considerable purity. Work by Brozek & Taylor (20) provides material to be integrated with that of Fleishmann. These authors seem more inclined than Fleishmann to emphasize the narrowness of the factors identified.

Of greatest importance is the finding in both programs that each test has

different meanings in different groups, a point consistent with the results of Heim and Zimmerman. Early and late in practice, tests and criterion skills have different makeup. In the Complex Coordination Test (62), only coordination is a major factor at all stages of practice. At early stages, cognitive abilities contribute heavily, but these diminish in influence while rate of movement and a test-specific factor become more prominent. The Discrimination Reaction Time Test behaves similarly.

Failure of assessment using data other than worksamples results in part from specifics in the criterion. When a person performs in a single situation for a long time, patterns of motor, cognitive, affective, and interpersonal behaviors learned specifically in that situation account for considerable variance [cf. Ferguson (55)]. Tests measure where a person starts from and predict how rapidly he will learn, not what reaction he will fix upon; yet that is ultimately required for assessment.

#### PERSONALITY MEASUREMENTS

Burt (23), Loevinger (114), Hutt (93), and Zubin (182, 183, 184) evaluate approaches to personality testing in important papers. Cattell and associates recount further factor analytic forays into personality (30, 31, 32). Their results seem too much open to sampling error, when correlations for 100 persons are made to yield 20 or so factors.

*Questionnaires: (a) Procedural studies.*—There is much debate regarding the merits of forced and unforced responses. Forced choice controls social desirability but cannot prevent all distortion [Durnall (46)]. Brogden (17) provides a technique for eliminating distortion when just one criterion is to be predicted and demonstrates its superior validity. Osburn and others (138) cite four recent unpublished studies claiming superior validity for forced-choice self-descriptions against leadership criteria.

Taken together with the past literature, and supplemented by three new sets of data (34, 36, 97), the Osburn study permits us to resolve contradictory findings on the forced versus unforced issue. A forced-choice item is less valid but less correlated with other items than is a comparable unforced item. Since each response in forced-choice is more saturated with valid variance, forcing is advantageous but only in a lengthy instrument. At lengths where the empirical difference is slight, the unforced form appears preferable because of its ease of construction and acceptability to subjects and its suitability for certain statistical analyses. In the unforced method, façade or response-set variance should be scored as such and often partialled out.

Fricke (68) found that a configural key based on patterns of responses consistently correlated higher with college achievement than a conventional key based on separate items. Secord (156) exposes serious faults in McQuitty's configural studies. Response sets such as acquiescence continue to appear as major sources of variance in questionnaires (31, 130).

*Questionnaires: (b) Specific instruments.*—The Edwards Personal Prefer-



ence Schedule (47) uses 225 paired comparisons to obtain self-reports on 15 Murray needs. The self-description is perhaps the most useful available among present tests for counseling. Façade is well removed; a social desirability score correlates only  $-.32$  to  $.32$  with Edwards scores, compared to  $.60$  with Guilford-Martin and reflected Taylor scores (see also 135).

The Taylor Anxiety (A) Scale has been a focus of research. Taking note of the finding that façade accounts for much of the variance, Kabrick (97) developed a refined forced-choice version. This scale has reliability  $.56$ , compared to  $.85$  for the A scale, and much less correlation with the MMPI K scale. A study of validity with a rather weak design leads to ambiguous but promising results. The Taylor scale correlates  $.92$  with MMPI Pt (16, 48), and both of these correlate around  $.90$  with a clinical rating of optimism (48).

Sampson & Bindra (148) report a zero correlation of A scores with clinical ratings while Buss and others (25) find a correlation of  $.60$ , but the latter data were obtained under conditions which might give an artificially high value. On the whole, the Taylor scale appears neither better nor worse than other adjustment questionnaires.

A number of investigators (3, 64, 106, 131, 153, 154, 158) try to determine whether different scores supposed to reflect authoritarianism or rigidity or both are correlated. The results are contradictory. In those studies which reject a general factor, the reliability of the measures is uncertain; in those which find positive relations, test-taking attitudes may provide an adequate explanation. Two studies report factor analysis of items. A general factor plus a group factor for religious items seems adequate to account for the O'Neil-Levinson data (137) but Saunders' (151) data require five factors representing different attitudes toward others.

Black (15) discusses use of MMPI with college women, reporting that their mean is substantially higher than the standardization group (see also 33, 71), describing cross-sectional trends with age, and summarizing relations of MMPI scores to peer ratings on 15 traits. These data make evident the necessity for determining the meaning of a particular score or pattern within any subculture, as opposed to applying one set of interpretations universally.

In an important program of research, Clark and others (34) have developed a Navy interest inventory with keys for Navy and civilian occupations. The scales separate men-in-job from men-in-general with about 35 per cent overlap, and correlate about  $.30$  with grades in technical school. Carter (27) finds a zero correlation of Kuder scores with grades in engineering. Lipsett & Wilson (111) followed 700 adults for one year, obtaining estimates of job satisfaction. The statistical treatment is unsatisfactory, but the data indicate higher satisfaction when people are in jobs consistent with measured ability and interest. Layton (103) reviews several current studies of the Strong Blank.

*Projective techniques: (a) Picture interpretations.*—Bellak (9) provides a

thoughtful but incautious manual on TAT and Children's Apperception Test. He regards the performance as a sample of ego-functioning rather than a sheer expression of drives. He gives particularly clear examples of interpretative procedure and rationale. More introductory and much less daring is Stein's revised TAT manual (161). He stays close to Murray's original views, in contrast to Bellak's psychoanalytic orientation. Both writers neglect validation.

That validation research is needed is demonstrated by Meyer & Tolman (127), who try to determine whether perceptions of TAT figures are representative of the subjects' attitudes toward significant others. Images of parent figures did not correspond to attitudes expressed in early therapeutic interviews. McIntyre (123) required subjects to describe, on MMPI items, themselves and also the TAT figures. One may quarrel with the assumption that similarity of these descriptions is a measure of projection. There was little projection in this sense, and subjects did not project more onto figures physically similar to themselves or onto vague figures. According to Milam (128), an unfriendly examiner elicits longer and more anxious TAT stories. Four other studies show effects of immediate motivational factors on thematic tests (66, 109, 134, 140).

Evidence favorable to Murray's approach comes from a sophisticated study by McArthur (122). Anthropological hypotheses from F. Kluckhohn regarding upper-class and middle-class values were found to be consistent with TAT differences. Among institutionalized delinquent boys, those with more fantasy aggression indulge in more overt aggression, if punishment press is not high [Mussen & Naylor (133)]. Thus TAT does predict behavior in this acting-out group. But in leaderless groups of college students, fantasy aggression correlated with overt aggression for authoritarians (who may be regarded as fearful) and not for equalitarians (Goodrich (70)). Several correlations higher than .30 were obtained for TAT "signs of anxiety" with ratings of assessors who had projective data available [Lindzey & Newburg (110)]. Carr (26) scored character of affect expressed, determining whether different projectives implied the same emotion. Overt TAT responses agree better with sentence completions, thematic TAT scores correspond better to Rorschach, and Rorschach has little in common with sentence completion.

An unsatisfactory procedure is used by Little & Schneidman (113) to "validate" TAT [also MMPI (112)]. Twenty-nine judges provided a criterion by *Q*-sorting after studying case material on a single patient. Seventeen other judges *Q*-sorted on the basis of TAT protocols; the average correlation with the composite criterion was .60. Such correlations cannot be interpreted as validity coefficients. To validate the test as a differential instrument it must be shown that with several superficially similar cases the correlation of the sort for each man with his own criterion is consistently higher than the cross-correlations. This MMPI study has additional faults. Good procedure is shown in the highly successful attempt by Horwitz & Cartwright (89) to validate specific hypotheses about a thematic test of group behavior.

De Charms and others (43) test whether a simple questionnaire measures need achievement in the same way as McClelland's projective tests. A value for achievement score, obtained from a questionnaire, elicits different facts. The value for achievement score measures tendency to be influenced by expert opinion. Since "Yes" is the scored answer to all questions, acquiescence is a component of the score. The research provides no definitive test whether a well-designed questionnaire could duplicate projective results. Ancona (2) shows that a negativistic, critical response set is significantly associated with need for achievement.

There is little evidence that the TAT protocol, examined for formal characteristics or process variables, is of more diagnostic value than any other sample of verbal behavior. The test produces meaningful thematic information, but much more research on interpretation is required. Considerable caution is required in interpreting thematic material as a random sample of perceptions and fantasies.

*Projective techniques:* (b) *Rorschach*.—Sarason (150) offers the first major report on Rorschach which simultaneously adheres to the standards of scientific psychology and reflects faithfully the clinical use of the test. He interprets the test as purposeful problem solving; how the subject interprets both test and clinician dictates the criteria he sets for his performance and the stress or motivation he feels. Sarason collates the literature, including unpublished theses.

Validation is competently studied in programs under Sarason (94, 149, 155; see 150), Phillips (102, 115, 129, 141), and Hamlin (12, 42, 76, 136). In the Hamlin program the basic question is whether the test is best interpreted piecemeal or globally. Cummings (42) designed an adjustment scale for rating response to each card. Psychiatrists ranked adjustment on the basis of case histories; these correlated .40 to .50 with Rorschach ratings, in a wide-range sample. When Newton (136) used an identical method save that Rorschach judgments were based on the protocol as a whole, validity was negligible. Hamlin (76), considering results of his entire program, believes that the ideal unit for judging is neither the single response nor the entire protocol, but is intermediate between them.

Rather than review separate studies in inadequate space, we shall attempt to summarize the present status of Rorschach hypotheses. Scientific study of the Rorschach has barely begun but is now evidenced in studies which translate clinical theory intelligently into experimental designs and adopt proper methodology. A remarkable proportion of these attempts are successful in rejecting the null hypothesis, but failures of Rorschach hypotheses and low-level relations are frequent enough to demand revision of accepted interpretations. The following impressions are based on less-than-exhaustive consideration of the current scene, drawing heavily on Sarason and on literature of prior years along with the new studies cited.

Number of responses has uncertain significance (101) but should be controlled in making other interpretations. Approach is a meaningful vari-

able indicating the subject's way of defining the task, but scoring should distinguish effortful, integrated W's from others [Harrington (78); Lane (102); Lofchie (115); Misch (129); Phillips & Framo (141)]. Character of Dd is more important than the number. S has some relation to opposition. Form dominance as represented in M, F, FC, and other scores, compared to vague and nonform-dominant responses, reflects effort to use outer controls, and form quality reflects ability to do so [Janoff (94); Sarason (149)].

Determinant scores require much further study. No confidence can be placed in alleged color-affect relations [Baughman (8); Keehn (98); Lazarus & Oldfield (104)], and shading-anxiety relations may disappear when vagueness and form-dominance are controlled [Cox & Sarason (37)]. "Shock" phenomena may represent simply the intellectual difficulty of dealing with certain cards. There is no support for usual interpretations of FM, m, C', or separate types of shading. Hypotheses about movement or *Erlebnistyp* have hitherto been entirely unconfirmed; several studies now support inconclusively an interpretation of M as inner-directedness and introspectiveness, but not of creativeness or adjustment [Barron (6); Mann (119); Schumer (155); Singer (160)]. While those who have rejected process scoring of the Rorschach appear to be in the wrong, this does not mean that the Rorschach "is validated." It is not demonstrated that the test is precise enough or invariant enough for clinical decisions. The test has repeatedly failed as a predictor of practical criteria. For those attributes which the Rorschach taps, better projective or objective tests can no doubt be built. The Rorschach's legitimate place appears to be as a wide-band (39) instrument singling out salient aspects of the personality regarding which more data must be obtained. There is nothing in the literature to encourage reliance on Rorschach interpretations.

*Projective techniques: (c) Other.*—A manual on a Sentence Completion Blank is offered by Holsopple & Miale (85); in this form, the test is a standardized open-end interview. Rather unfavorable evidence on the Picture-Frustration Test is reported by Lindzey & Goldwyn (108) and Vane (170).

Considering all work on projectives, one is impressed by the vastly improved quality of the best available research over the best of a few years back. The most bizarre projectives are fading away, the nomothetic studies of Rorschach are forming into a coherent whole, and the TAT studies are becoming analytic and more informative. One notices, however, the absence of worth-while research on idiographic interpretation, which the clinical writers advocate as the most fruitful manner of using projectives.

#### CRITERION-ORIENTED STUDIES: PERSONALITY

To many testers, a most notable event was the publication in *Fortune* of a sharp criticism of the use of personality tests for selecting high-level employees (177). Although Whyte's criticisms are less severe than some remarks in the Buros yearbooks, the appearance of such statements in a public place disturbed many applied psychologists. The article gave insufficient

credit to those industrial psychologists who perform careful validations where they use inventories, but the fact remains that inventories of negligible validity, which penalize the person who departs from a conventional and placid self-description, are being widely applied.

Hollander (83) reviews studies of the buddy rating technique and concludes that these procedures have substantial validity as predictors of Officer Candidate School performance and similar criteria. Several improvements are suggested.

The most successful situational procedure is the leaderless group discussion. Arbous (4) reports validity as high as .60 for executive selection. Lévy-Leboyer (107) and Bass (7) review a large amount of research which shows favorable validities, analyze the nature of the test variance (see also 28), and suggest precautions in interpretation. Both of these reviews are nice examples of construct validation.

Fiske (58) points out that situation tests are useful chiefly where capacities account for little of the criterion variance. Weislogel & Schwartz (175) stress the need to elicit typical performance, to score behavior on a frequency or presence-absence basis, and to interpret records in terms of probable job success. Strangely missing from such discussions is the fact that, as was noted in connection with abilities, while "typical habits" determine how a person acts when he enters a criterion situation, he soon forms new reaction patterns which account for much variance. It is on this rock that situational measures founder unless the relevant habits are highly generalized. A similar comment applies to objective tests of temperament. The latter are omitted from this review because no encouraging studies were found.

Questionnaires have been successful in postdicting delinquency. Hanley (77) took items from the California Personality Inventory which discriminated Navy men confined for violations. Substantial separation was achieved in cross-validation. Gough (72) found that the California Psychological Inventory discriminated very adequately between delinquent and nondelinquent adults. Another remarkable empirical program is the Brown-Holtzman work on prediction of overachievement (18, 19, 86). Although this goal has eluded hundreds of investigators, with 75 Likert-type items on study attitudes they were able to obtain correlations of .50 with grades and multiple *R*'s of .70. The key has sufficient generality to work in many, but not all, schools. Barron (5) discusses correlates of success in graduate school, but cross-validation of these results is required.

It is an open question whether empirical MMPI scales have practical value. A previously reported scale for identifying schizophrenics is not able to differentiate adequately between these and other patients (143, 146). A scale for "sexual deviation" is made to differentiate imprisoned sex offenders from male education majors in college (120). But such a differentiation must be confounded with many variables other than sexual, and the offenders are not representative of sex deviants in general. Della Piana & Gage (45) show

that while the Minnesota Teacher Attitude Inventory predicts acceptance of teachers in classes where pupils care little about accomplishment, the test has no predictive value where pupils want to learn. Two discouraging empirical studies are reported: Tyler's attempt (169) to predict teaching success from inventories, and French's (67) on leadership. While excellent theory behind item writing may permit one to develop an empirical key on a small sample, a study without theory requires huge samples for any hope of success. Papers suggesting basic limitations of empirical use of questionnaires are presented by Moran *et al.* (132) and Rosen (144). The former show that readily available facts such as location of residence predict response to hospitalization as validly as questionnaires usually predict. Rosen shows that a valid instrument may have no value where any decision the predictor permits involves an intolerable false positive rate.

An excellent critique by Yates (180) on procedures for assessing brain damage concludes with these remarks of very general application:

A purely empirical approach is unlikely to yield satisfactory results, nor is an approach based on a theory which has not been adequately tested experimentally. A satisfactory test of brain damage should be based on a reasonable theory . . . supported by adequate statistical treatment, taking . . . into account all relevant variables. Such an approach would at least help to overcome the impasse which seems to have been reached with many of the tests reviewed. . . . While most workers in this field report satisfactory discrimination, the constant "dog eat dog" method by which one set of signs is set aside as unusable and replaced by a new set by subsequent workers does not inspire confidence in the most recent method, that of Dörken and Kral.

Sure enough, the first careful cross-validation (57) eliminates Dörken and Kral from consideration. The Piotrowski system performs with fair validity, miscalling two nonorganics out of 34, and 52 organics out of 84. This validity is encouraging, since the investigators required their instrument to discriminate organics from other neurologic patients, not from college students or other irrelevant controls.

#### CONCLUSION

Assessment research involves two perspectives, the immediate and the remote. Applied psychologists have to make decisions every day about individuals, and they cannot wait for better techniques. As researchers, however, their concern is with clarifying the nature of man and the methods of investigating him, with practical techniques only an eventual aim. The techniques and studies which count for most in one perspective mean the least in the other.

For today's applied psychologist, minimum inference is clearly best. The test which duplicates the general factor underlying the day-to-day demands of the criterion situation is a good predictor. So is a patiently developed empirical key using diverse predictors. Where such minimum-inference methods are unavailable or too costly, inferences from test to trait to pre-

dicted criterion behavior are useful provided the inferences are checked immediately and at little cost. This situation will change very slowly. Factor research is unlikely to disclose key dimensions that unlock all doors. Factors are many and hard to measure reliably. Experimental methods are showing that far more than the "content" of a test accounts for variance in score. Specific reactions developed in the criterion situation may remain stubbornly unpredictable.

Ultimate progress depends on advances in psychology, not on test engineering. Sturdy links are being forged to mediate inferences from tests to construct interpretations. But the critical inference from knowledge about the individual to practical prediction, i.e., about the relation of constructs to criterion performance, is being ignored. Harrower (79) tells us that she failed as an assessor in the Holtzman-Sells study, and her reason is illuminating. Some of the men who succeeded had test records so abnormal that she cannot believe these men functioned as good pilots. But they did. Until we understand how men function in each criterion situation, methods of assessment and guidance which demand inference cannot succeed.

## LITERATURE CITED

1. Anastasi, A., "The Measurement of Abilities," *J. Counseling Psychol.*, **1**, 164-68 (1954)
2. Ancona, L., "La C. D. motivazione al successo (need for achievement) in termini di 'response set' di acquiescenza e di negativismo," *Arch. psicol. neurol. psychiat.*, **15**, 158-66 (1954)
3. Applezweig, D. G., "Some Determinants of Behavioral Rigidity," *J. Abnormal Social Psychol.*, **49**, 224-28 (1954)
4. Arbous, A. G., *Selection for Industrial Leadership* (Oxford University Press, London, England, 180 pp., 1954)
5. Barron, F., *Personal Soundness in University Graduate Students* (University of California Press, Berkeley, Calif., 32 pp., 1954)
6. Barron, F., "Threshold for the Perception of Human Movement in Inkblots," *J. Consulting Psychol.*, **19**, 33-38 (1955)
7. Bass, B. M., "The Leaderless Group Discussion," *Psychol. Bull.*, **51**, 465-92 (1954)
8. Baughman, E. E., "A Comparative Analysis of Rorschach Forms with Altered Stimulus Characteristics," *J. Projective Techniques*, **18**, 151-64 (1954)
9. Bellak, L., *The Thematic Apperception Test and the Children's Apperception Test in Clinical Use* (Grune & Stratton, Inc., New York, N. Y., 282 pp., 1954)
10. Bennett, G. K., and Gelink, M., *Manual for Short Employment Tests* (Psychological Corporation, New York, N. Y., 11 pp., 1954)
11. Berdie, R. F., "Aptitude, Achievement, Interest, and Personality Tests: A Longitudinal Comparison," *J. Appl. Psychol.*, **39**, 103-14 (1955)
12. Bialick, I., and Hamlin, R. M., "The Clinician as Judge: Details of Procedure in Judging Projective Material," *J. Consulting Psychol.*, **18**, 239-42 (1954)
13. Biesheuvel, S., "The Measurement of Occupational Aptitudes in a Multi-Racial Society," *Occupational Psychol. (London)*, **28**, 189-96 (1954)
14. Biesheuvel, S., Jacobs, G. F., and Cowley, J. J., "Maladjustments of Military Personnel," *J. Natl. Inst. for Personnel Research*, **5**, 138-68 (1953)
15. Black, J. D., "The Use of the MMPI with Normal Persons," in *New Perspectives in Counseling*, 33-48 (Hewer, V. H., Ed., University of Minnesota Press, Minneapolis, Minn., 66 pp., 1955)
16. Brackbill, G., and Little, K. B., "MMPI Correlates of the Taylor Scale of Manifest Anxiety," *J. Consulting Psychol.*, **18**, 433-36 (1954)
17. Brogden, H. E., "A Rationale for Minimizing Distortion in Personality Questionnaire Keys," *Psychometrika*, **19**, 141-48 (1954)
18. Brown, W. F., and Holtzman, W. H., "A Study-Attitudes Questionnaire for Predicting Academic Success," *J. Educ. Psychol.*, **46**, 75-84 (1955)
19. Brown, W. F., and Holtzman, W. H., *Brown-Holtzman Survey of Study Habits and Attitudes* (Psychological Corporation, New York, N. Y., 8 pp., 1953)
20. Brozek, J., and Taylor, H. L., "Tests of Motor Functions in Investigations in Fitness," *Am. J. Psychol.*, **67**, 590-611 (1954)
21. Bruner, J. S., and Tagiuri, R., "The Perception of People," in *Handbook of Social Psychology*, 634-54 (Lindzey, G., Ed., Addison-Wesley Publishing Co., Inc., Cambridge, Mass., 2 vols., 1226 pp., 1954)
22. Burgemeister, B. B., Blum, L. H., and Lorge, I., *Columbia Mental Maturity Scale* (World Book Co., Yonkers, N. Y., 13 pp., 1954)
23. Burt, C., "The Assessment of Personality," *J. Mental Sci.*, **100**, 1-28 (1954)



24. Burt, C., "The Differentiation of Intellectual Ability," *Brit. J. Educ. Psychol.*, **24**, 76-90 (1954)
25. Buss, A. H., Wiener, M., Durkee, A., and Baer, M., "The Measurement of Anxiety in Clinical Situations," *J. Consulting Psychol.*, **19**, 125-29 (1955)
26. Carr, A. C., "Intra-Individual Consistency in Response to Tests of Varying Degrees of Ambiguity," *J. Consulting Psychol.*, **18**, 251-58 (1954)
27. Carter, G. C., "Kuder Preference Record Scores and Success in Engineering College," *J. Counseling Psychol.*, **1**, 196 (1954)
28. Carter, L. F., "Evaluating the Performance of Individuals as Members of Small Groups," *Personnel Psychol.*, **7**, 477-84 (1954)
29. Cattell, R. B., "Growing Points in Factor Analysis," *Australian J. Psychol.*, **6**, 105-40 (1954)
30. Cattell, R. B., Dubin, S. S., and Saunders, D. R., "Personality Structure in Psychotics by Factorization of Objective Clinical Tests," *J. Mental Sci.*, **100**, 154-76 (1954)
31. Cattell, R. B., Dubin, S. S., and Saunders, D. R., "Verification of Hypothesized Factors in One Hundred and Fifteen Objective Personality Test Designs," *Psychometrika*, **19**, 209-30 (1954)
32. Cattell, R. B., and Saunders, D. R., "Beiträge zur Faktoren-Analyse der Persönlichkeit," *Z. Exptl. Angew. Psychol.*, **2**, 325-57 (1954)
33. Clark, J. H., "The Interpretation of the MMPI Profiles of College Students: Mean Scores for Male and Female Groups," *J. Social Psychol.*, **40**, 319-21 (1954)
34. Clark, K. E., Gee, H. H., and Perry, D., *Measurement of Interest Patterns* (University of Minnesota, Minneapolis, Minn., 46 pp., 1953)
35. Clarke, A. D. B., and Clarke, A. M., "How Constant Is the I. Q.?", *Lancet*, **II**, 877-85 (1953)
36. Coombs, C. H., Milholland, J. E., and Womer, F. B., *The Assessment of Partial Knowledge in Objective Testing* (Engineering Research Institute, University of Michigan, Ann Arbor, Mich., 55 pp., 1955)
37. Cox, F. N., and Sarason, S. B., "Test Anxiety and Rorschach Performance," *J. Abnormal Social Psychol.*, **49**, 371-77 (1954)
38. Cronbach, L. J., "Report on a Psychometric Mission to Clinicia," *Psychometrika*, **19**, 263-70 (1954)
39. Cronbach, L. J., "The Counselor's Problems from the Perspective of Communication Theory," in *New Perspectives in Counseling*, 3-19 (Hewer, V. H., Ed., University of Minnesota Press, Minneapolis, Minn., 66 pp., 1955)
40. Cronbach, L. J., and Gleser, G. C., "Review of Stephenson, W., 'The Study of Behavior'," *Psychometrika*, **19**, 327-33 (1954)
41. Crow, W. J., *A Methodological Study of Social Perceptiveness* (Doctoral thesis, Univ. of Colorado, Boulder, Colo., 139 pp., 1954)
42. Cummings, S. T., "The Clinician as Judge: Judgments of Adjustment from Rorschach Single-Card Performance," *J. Consulting Psychol.*, **18**, 243-47 (1954)
43. De Charms, R., Morrison, H. W., and Reitman, W., *Behavioral Correlates of Directly and Indirectly Measured Achievement Motivation* (Wesleyan University, Middletown, Conn., 13 pp., 1954)
44. de Groot, A. D., "Scientific Personality Diagnosis," *Acta Psychol.*, **10**, 220-41 (1954)

45. Della Piana, G. M., and Gage, N. L., "Pupils' Values and the Validity of the Minnesota Teacher Attitude Inventory," *J. Educ. Psychol.*, **46**, 167-78 (1955)
46. Durnall, E. J., Jr., "Falsification of Interest Patterns on the Kuder Preference Record," *J. Educ. Psychol.*, **45**, 240-43 (1954)
47. Edwards, A. L., *Edwards Personal Preference Schedule* (Psychological Corporation, New York, N. Y., 18 pp., 1953)
48. Eriksen, C. W., and Davids, A., "The Meaning and Clinical Validity of the Taylor Anxiety Scale and the Hysteria-Psychasthenia Scales from the MMPI," *J. Abnormal Social Psychol.*, **50**, 135-37 (1955)
49. Eysenck, H. J., "A Dynamic Theory of Anxiety and Hysteria," *J. Mental Sci.*, **101**, 28-51 (1955)
50. Eysenck, H. J., "Abord statistique et experimental du probleme typologique dans la personnalité nevrotique, psychotique et normale," *L'Évolution Psychiat.*, **3**, 377-404 (1954)
51. Eysenck, H. J., "Probleme der diagnostischen Untersuchung und Demonstration des Charakter-Interpretationstestes," *Z. Exptl. Angew. Psychol.*, **2**, 1-32 (1954)
52. Eysenck, H. J., "Review of Stephenson, W., 'The Study of Behavior'," *J. Educ. Psychol.*, **45**, 374-76 (1954)
53. Eysenck, H. J., "Zur Theorie der Persönlichkeitsmessung," *Psychol. u. Persönlichkeitsforsch.*, **2**, 87-101, 171-87 (1954)
54. Fahmy, M., *Initial Exploring of the Shilluk Intelligence* (Dar Misr Printing House, 37A Kamel Sidky St., Cairo, Egypt, 32 pp., 1954)
55. Ferguson, G. A., "On Learning and Human Ability," *Can. J. Psychol.*, **8**, 95-112 (1954)
56. Fisher, S., and Fisher, R., "A Study of the Relationship between Personal Insecurity and One's Theoretical Orientation toward Psychological Methodology," *Am. Psychologist*, **9**, 368 (1954)
57. Fisher, J., Gonda, T. A., and Little, K. B., "The Rorschach and Central Nervous System Pathology: a Cross-Validation Study," *Am. J. Psychiat.*, **7**, 487-92 (1954)
58. Fiske, D. W., "Why Do We Use Situational Performance Tests?," *Personnel Psychol.*, **7**, 464-69 (1954)
59. Flanagan, J. C., *Flanagan Aptitude Classification Tests* (Science Research Associates, Chicago, Ill., 135 pp., 1953)
60. Fleishman, E. A., "Dimensional Analysis of Psychomotor Abilities," *J. Exptl. Psychol.*, **48**, 437-54 (1954)
61. Fleishman, E. A., and Hempel, W. E., Jr., "A Factor Analysis of Dexterity Tests," *Personnel Psychol.*, **7**, 15-32 (1954)
62. Fleishman, E. A., and Hempel, W. E., Jr., "Changes in Factor Structure of a Complex Psychomotor Test as a Function of Practice," *Psychometrika*, **19**, 239-52 (1954)
63. Fleishman, E. A., and Hempel, W. E., Jr., "Factorial Analysis of Complex Psychomotor Performance," *Research Bull. No. 54-12* (Personnel and Training Research Center, Lackland Air Force Base, San Antonio, Texas, 1954)
64. Forster, N. C., Vinacke, W. E., and Digman, J. M., "Flexibility and Rigidity in a Variety of Problem Situations," *J. Abnormal Social Psychol.*, **50**, 211-16 (1955)
65. Franks, C., *An Experimental Study of Conditioning as Related to Mental Abnormality* (Doctoral thesis, Univ. of London, London, England, 1954)

66. French, E. G., "Experimental Investigation of a Method of Measuring Achievement Motivation," *Am. Psychologist*, **9**, 371 (1954)
67. French, J. W., "The Validity of Some Objective Personality Tests for a Leadership Criterion," *Educ. Psychol. Measurement*, **14**, 34-49 (1954)
68. Fricke, B. G., *The Development of an Empirically Validated Personality Test Employing Configural Analysis for the Prediction of Academic Achievement* (Doctoral thesis, Univ. of Minnesota, Minneapolis, Minn., 1954)
69. Gleser, G. C., Haddock, J., Starr, P., and Ulett, G. A., *Psychiatric Screening of Flying Personnel: Interrater Agreement on the Basis of Psychiatric Interviews* (School of Aviation Medicine, Randolph Field, Texas, 7 pp., 1954)
70. Goodrich, D. C., "Aggression in the Projective Tests and Group Behavior of Authoritarian and Equalitarian Subjects," *Am. Psychologist*, **9**, 380 (1954)
71. Goodstein, L. D., "Regional Differences in MMPI Responses Among Male College Students," *J. Consulting Psychol.*, **18**, 437-41 (1954)
72. Gough, H. G., "Systematic Validation of a Test for Delinquency," *Am. Psychologist*, **9**, 381 (1954)
73. Guttman, L., "A New Approach to Factor Analysis: the Radex," in *Mathematical Thinking in the Social Sciences*, 258-348 (Lazarsfeld, P. F., Ed., The Free Press, Glencoe, Ill., 444 pp., 1954)
74. Haggard, E. A., "Social Status and Intelligence: An Experimental Study of Certain Cultural Determinants of Measured Intelligence," *Genet. Psychol. Monographs*, **49**, 141-86 (1954)
75. Halbower, C. C., *A Comparison of Actuarial versus Clinical Prediction to Classes Discriminated by MMPI* (Doctoral thesis, Univ. of Minnesota, Minneapolis, Minn., 1955)
76. Hamlin, R. M., "The Clinician as Judge: Implications of a Series of Studies," *J. Consulting Psychol.*, **18**, 233-38 (1954)
77. Hanley, C., *An Inventory of Personal Opinions* (U. S. Naval Retraining Command, San Diego, Calif., 17 pp., 1954)
78. Harrington, R. W., "Maladaptive Responses to Frustration Predicted from Rorschach Color Responses," *J. Consulting Psychol.*, **18**, 455-58 (1954)
79. Harrower, M., "Clinical Aspects of Failures in the Projective Techniques," *J. Projective Techniques*, **18**, 294-302 (1954)
80. Heim, A. W., "Les conventions de base des tests d'intelligence," *Enfance*, **7**, 105-12 (1954)
81. Hempel, W. E., Jr., and Fleishman, E. A., "A Factor Analysis of Physical Proficiency and Manipulative Skill," *J. Appl. Psychol.*, **39**, 12-16 (1955)
82. Hofstaetter, P. R., "The Changing Composition of 'Intelligence'; A Study in T-Technique," *J. Genet. Psychol.*, **85**, 159-64 (1954)
83. Hollander, E. P., "Buddy Ratings: Military Research and Industrial Implications," *Personnel Psychol.*, **7**, 385-93 (1954)
84. Holloway, H. D., "Effects of Training on the SRA Primary Mental Abilities (Primary) and the WISC," *Child Development*, **25**, 253-63 (1954)
85. Holsopple, J. Q., and Miale, E. R., *Sentence Completion, A Projective Method for the Study of Personality* (Charles C Thomas, Springfield, Ill., 177 pp., 1954)
86. Holtzman, W. H., Brown, W. F., and Farquhar, W. G., "The Survey of Study Habits and Attitudes: a New Instrument for the Prediction of Academic Success," *Educ. Psychol. Measurement*, **14**, 726-32 (1954)

87. Holtzman, W. H., and Sells, S. B., "Prediction of Flying Success by Clinical Analysis of Test Protocols," *J. Abnormal Social Psychol.*, **49**, 485-90 (1954)
88. Holzinger, K. J., and Crowder, N. A., *Holzinger-Crowder Uni-Factor Tests* (World Book Company, Yonkers, N. Y., 28 pp., 1955)
89. Horwitz, M., and Cartwright, D., "A Projective Method for the Diagnosis of Group Properties," *Human Relations*, **6**, 397-410 (1953)
90. Husén, T., "La validité des interviews par rapport à l'âge, au sexe et à la formation des interviewés," *Trav. Humain*, **17**, 60-67 (1954)
91. Husén, T., "Über die Begabung von Zwillingen," *Psychol. Beiträge*, **1**, 137-45 (1953)
92. Hurtig, M., "Recherche sur la perfectibilité," *Enfance*, **7**, 317-26 (1954)
93. Hutt, M. L., "Toward an Understanding of Projective Testing," *J. Projective Techniques*, **18**, 197-201 (1954)
94. Janoff, I. Z., *The Relation between Rorschach Form Quality Measures and Children's Behavior* (Doctoral thesis, Yale Univ., New Haven, Conn., 1951)
95. Jensen, M. B., and Schmid, J., "An Analysis of Some Clinical Judgments on Male Basic Airmen Who Failed the Group Psychological Tests," *J. Clin. Psychol.*, **10**, 325-32 (1954)
96. Jones, E. E., "Authoritarianism as a Determinant of First-Impression Formation," *J. Personality*, **23**, 107-27 (1954)
97. Kabrick, R. P., *Predictive Value of a Revised Forced Choice Form of the Manifest Anxiety Scale* (Iowa State University, Iowa City, Iowa, 25 pp., 1954)
98. Keehn, J. D., "A Re-interpretation of the Role Played by Color in the Rorschach Test," *Brit. J. Med. Psychol.*, **27**, 89-93 (1954)
99. Kelly, E. L., "The Place of Situation Tests in Evaluating Clinical Psychologists," *Personnel Psychol.*, **7**, 484-92 (1954)
100. Kelly, G. A., *The Psychology of Personal Constructs. I. A Theory of Personality* (W. W. Norton and Co., New York, N. Y., 556 pp., 1955)
101. Kornetsky, C., and Gerard, D. L., "Effect of Increasing the Number of Rorschach Responses on Sum C and M: A Note on Fiske and Baughman's Study," *J. Abnormal Social Psychol.*, **49**, 592-93 (1954)
102. Lane, J. R., "Social Effectiveness and Developmental Level," *J. Personality*, **23**, 274-83 (1955)
103. Layton, W. L., "Theory and Research on the Strong Vocational Interest Blank: A Conference Report," *J. Counseling Psychol.*, **2**, 10-12 (1955)
104. Lazarus, R. S., and Oldfield, M., "Rorschach Responses and the Influence of Color," *J. Personality*, **23**, 356-72 (1955)
105. Léon, A., "Quelques remarques sur les apports des dernières rencontres de psychologues," *Enfance*, **7**, 381-92 (1954)
106. Levitt, E. E., "Studies in Intolerance of Ambiguity: I. The Decision-Location Test with Grade School Children," *Child Development*, **24**, 263-68 (1953)
107. Lévy-Leboyer, C., "Les tests de groupe," *Trav. Humain*, **18**, 17-39 (1955)
108. Lindzey, G., and Goldwyn, R. M., "Validity of the Rosenzweig Picture-Frustration Study," *J. Personality*, **22**, 519-47 (1954)
109. Lindzey, G., and Herman, P. S., "Thematic Apperception Test: A Note on Reliability and Situational Validity," *J. Projective Techniques*, **19**, 36-42 (1955)
110. Lindzey, G., and Newburg, A. S., "Thematic Apperception Test: A Tentative Appraisal of Some 'Signs' of Anxiety," *J. Consulting Psychol.*, **18**, 389-95 (1954)

111. Lipsett, L., and Wilson, J. W., "Do 'Suitable' Interests and Mental Ability Lead to Job Satisfaction?," *Educ. Psychol. Measurement*, **14**, 373-80 (1954)
112. Little, K. B., and Shneidman, E. S., "The Validity of MMPI Interpretations," *J. Consulting Psychol.*, **18**, 425-28 (1954)
113. Little, K. B., and Shneidman, E. S., "The Validity of Thematic Projective Technique Interpretations," *J. Personality*, **23**, 285-94 (1955)
114. Loevinger, J., "Some Principles of Personality Measurement," *Educ. Psychol. Measurement*, **15**, 3-17 (1955)
115. Lofchie, S. H., "The Performance of Adults under Distraction Stress: A Developmental Approach," *J. Psychol.*, **39**, 109-16 (1955)
116. Lorge, I., and Thorndike, R. L., *Lorge-Thorndike Intelligence Tests* (Houghton Mifflin Co., Boston, Mass., 9 pp., 1955)
117. Luborsky, L., "Selecting Psychiatric Residents: Survey of the Topeka Research," *Bull. Menninger Clinic*, **18**, 252-59 (1954)
118. Mackie, R. R., Strayer, F. K., and Buckner, D. N., *Validation of the U. S. Medical Research Laboratory Personnel Assessments* (Management and Marketing Research Corp., Los Angeles, Calif., 43 pp., 1954)
119. Mann, L., *The Relation of Rorschach Indices of Extratension and Introversion to a Measure of Responsiveness to the Immediate Environment* (Doctoral thesis, Univ. of North Carolina, Chapel Hill, N. C., 1953)
120. Marsh, J. T., Hilliard, J., and Liechti, R., "A Sexual Deviation Scale for the MMPI," *J. Consulting Psychol.*, **19**, 55-59 (1955)
121. McArthur, C., "Analyzing the Clinical Process," *J. Counseling Psychol.*, **1**, 203-7 (1954)
122. McArthur, C., "Personality Differences between Middle and Upper Classes," *J. Abnormal Social Psychol.*, **50**, 247-54 (1955)
123. McIntyre, C. J., "Sex, Age, and Iconicity as Factors in Projective Film Tests," *J. Consulting Psychol.*, **18**, 337-43 (1954)
124. Meehl, P. E., *Clinical vs. Statistical Prediction* (Univ. of Minnesota Press, Minneapolis, Minn., 149 pp., 1954)
125. Meili, R., "Ein Zerrbild der Test-Psychologie," *Psychol. Rundschau*, **5**, 81-86 (1954)
126. Melton, R. S., *A Comparison of Clinical and Actuarial Methods of Prediction with an Assessment of the Relative Accuracy of Different Clinicians* (Doctoral thesis, Univ. of Minnesota, Minneapolis, Minn., 124 pp., 1952)
127. Meyer, M. M., and Tolman, R. S., "Correspondence between Attitudes and Images of Parental Figures in TAT stories and in Therapeutic Interviews," *J. Consulting Psychol.*, **19**, 79-82 (1955)
128. Milam, J. R., "Examiner Influences on Thematic Apperception Test Stories," *J. Projective Techniques*, **18**, 221-26 (1954)
129. Misch, R. C., *The Relationship of Motoric Inhibition to Developmental Level and Ideational Functioning* (Doctoral thesis, Clark Univ., Worcester, Mass., 1953)
130. Mitzel, H. E., Rabinowitz, W., and Ostriecher, L. M., *Effect of Certain Response Sets on Valid Test Variance* (Division of Teacher Education, College of the City of New York, New York, N. Y., 23 pp., 1955)
131. Mooney, C. M., "A Factorial Study of Closure," *Can. J. Psychol.*, **8**, 51-60 (1954)
132. Moran, L. J., Fairweather, G. W., Morton, R. B., and McGaughran, L. S., "The Use of Demographic Characteristics in Predicting Response to Hospitalization for Tuberculosis," *J. Consulting Psychol.*, **19**, 65-70 (1955)

133. Mussen, P. H., and Naylor, H. K., "The Relationships between Overt and Fantasy Aggression," *J. Abnormal Social Psychol.*, **49**, 235-40 (1954)
134. Mussen, P. H., and Scodel, A., "The Effects of Sexual Stimulation under Varying Conditions on TAT Sexual Responsiveness," *J. Consulting Psychol.*, **19**, 90-91 (1955)
135. Navran, L., and Stauffacher, J. C., "Social Desirability as a Factor in Edwards Personality Preference Schedule Performance," *J. Consulting Psychol.*, **18**, 442 (1954)
136. Newton, R. L., "The Clinician as Judge: Total Rorschachs and Clinical Case Material," *J. Consulting Psychol.*, **18**, 248-50 (1954)
137. O'Neil, W. M., and Levinson, D. J., "A Factorial Exploration of Authoritarianism and Some of its Ideological Concomitants," *J. Personality*, **22**, 449-63 (1954)
138. Osburn, H. G., Lubin, A., Loeffler, J. C., and Tye, V. M., "The Relative Validity of Forced Choice and Single Stimulus Self Description Items," *Educ. Psychol. Measurement*, **14**, 407-17 (1954)
139. Osgood, C. E., and Luria, Z., "A Blind Analysis of a Case of Multiple Personality Using the Semantic Differential," *J. Abnormal Social Psychol.*, **49**, 579-91 (1954)
140. Parrish, J., and Rethlingshafer, D., "A Study of the Need to Achieve in College Achievers and Non-Achievers," *J. Gen. Psychol.*, **50**, 209-26 (1954)
141. Phillips, L., and Framo, J. L., "Developmental Theory Applied to Normal and Psychopathological Perception," *J. Personality*, **22**, 464-74 (1954)
142. Porteus, S. D., "Maze Test Qualitative Aspects," *Brit. J. Med. Psychol.*, **27**, 72-79 (1954)
143. Quay, H., and Rowell, J. T., "The Validity of a Schizophrenic Screening Scale of the MMPI," *J. Clin. Psychol.*, **11**, 92-93 (1955)
144. Rosen, A., "Detection of Suicidal Patients: An Example of Some Limitations in the Prediction of Infrequent Events," *J. Consulting Psychol.*, **18**, 397-404 (1954)
145. Rosenwald, A. K., "An Assessment of the Current Role of Psychodiagnostic Testing," *J. Consulting Psychol.*, **18**, 311-15 (1954)
146. Rubin, H., "Validity of a Critical-Item Scale for Schizophrenia on the MMPI," *J. Consulting Psychol.*, **18**, 219-20 (1954)
147. Rulon, P. J., and Schweiker, R. F., *Validation of a Non-Verbal Test of Military Trainability. The Semantic Test of Intelligence* (Harvard University, Cambridge, Mass., 49 pp., 1953)
148. Sampson, H., and Bindra, D., "'Manifest' Anxiety, Neurotic Anxiety, and the Rate of Conditioning," *J. Abnormal Social Psychol.*, **49**, 256-59 (1954)
149. Sarason, E. K., *The Discriminatory Value of the Rorschach Test Between Two Etiologically Different, Mentally Defective Groups* (Doctoral thesis, Clark Univ., Worcester, Mass., 1950)
150. Sarason, S. B., *The Clinical Interaction, with Special Reference to the Rorschach* (Harper & Brothers, New York, N. Y., 425 pp., 1954)
151. Saunders, D. R., *A Factor Analysis of Some Personality Items in a Domain of Social Identification* (Educational Testing Service, Princeton, N. J., 15 pp., 1954)
152. Schafer, R., *Psychoanalytic Interpretation in Rorschach Testing* (Grune & Stratton, Inc., New York, N. Y., 446 pp., 1954)

153. Scheier, I. H., "An Evaluation of Rigidity Factors," *Can. J. Psychol.*, **8**, 157-63 (1954)
154. Schmidt, H. O., Fonda, C. P., and Wesley, E. L., "A Note on Consistency of Rigidity as a Personality Variable," *J. Consulting Psychol.*, **18**, 450 (1954)
155. Schumer, F. C., *Some Behavioral Correlates of Rorschach Human Movement Responses* (Doctoral thesis, Yale Univ., New Haven, Conn., 1949)
156. Secord, P. F., "'Personality Integration' in Responses to Self-Inventories," *J. Personality*, **23**, 308-16 (1955)
157. Segel, D., and Raskin, E., *Multiple Aptitude Tests* (California Test Bureau, Los Angeles, Calif., 1955)
158. Siegel, S., "Certain Determinants and Correlates of Authoritarianism," *Genet. Psychol. Monographs*, **49**, 187-229 (1954)
159. Simoneit, M., "Zur Kritik der Test-Psychologie," *Psychol. Rundschau*, **5**, 44-80 (1954)
160. Singer, J. L., and Herman, J., "Motor and Fantasy Correlates of Rorschach Human Movement," *J. Consulting Psychol.*, **18**, 325-31 (1954)
161. Stein, M. I., *The Thematic Apperception Test, Revised Edition* (Addison-Wesley Publishing Co., Inc., Cambridge, Mass., 365 pp., 1955)
162. Steiner, I. D., "Ethnocentrism and Tolerance of Trait 'Inconsistency'," *J. Abnormal Social Psychol.*, **49**, 349-54 (1954)
163. Stephenson, W., *The Study of Behavior* (University of Chicago Press, Chicago, Ill., 376 pp., 1953)
164. Symonds, P. M., "Are Projective Test Data Valid Bases for Prediction?," *J. Projective Techniques*, **18**, 515-19 (1954)
165. Terman, L. M., "The Discovery and Encouragement of Exceptional Talent," *Am. Psychologist*, **9**, 221-30 (1954)
166. Thurstone, L. L., and Thurstone, T. G., *Examiner Manual for the SRA Primary Mental Abilities, Ages 7 to 11*, 2nd ed. (Science Research Associates, Chicago, Ill., 21 pp., 1954)
167. Tiedeman, D. V., and Bryan, J. G., "Prediction of College Field of Concentration," *Harvard Educ. Rev.*, **24**, 122-39 (1954)
168. Tiffin, J., and Lawshe, C. H., *Examiner Manual for the Adaptability Test* (Science Research Associates, Chicago, Ill., 6 pp., 1954)
169. Tyler, F. T., *The Prediction of Student-Teaching Success from Personality Inventories* (University of California Press, Berkeley, Calif., 81 pp., 1954)
170. Vane, J. R., "Implications of the Performance of Delinquent Girls on the Rosenzweig Picture-Frustration Study," *J. Consulting Psychol.*, **18**, 414 (1954)
171. Vernon, P. E., "Symposium on the Effect of Coaching and Practice on Intelligence Tests. V. Conclusions," *Brit. J. Educ. Psychol.*, **24**, 57-63 (1954)
172. Walker, W. B., *An Investigation of the Effectiveness of Communication between Psychologists and Sales Executives Through Personnel Audit Reports* (Doctoral thesis, Western Reserve Univ., Cleveland, Ohio, 209 pp., 1955)
173. Wand, B., and Mollenkopf, W. G., *Selection and Classification Tests for Women* (Educational Testing Service, Princeton, N. J., 64 pp., 1954)
174. Wechsler, D., *Wechsler Adult Intelligence Scale Manual* (Psychological Corporation, New York, N. Y., 110 pp., 1955)
175. Weislogel, R. L., and Schwartz, P. A., "Some Practical and Theoretical Problems in Situational Testing," *Educ. Psychol. Measurement*, **15**, 39-46 (1955)

176. Wertheimer, M., "The Differential Satiability of Schizophrenic and Normal Subjects: A Test of a Deduction from the Theory of Figural After-Effects," *J. Gen. Psychol.*, **51**, 291-99 (1954)
177. Whyte, W. H., Jr., "The Fallacies of 'Personality' Testing," *Fortune*, **50**, 117-21 (1954)
178. Wilson, J. W., "Correlation of Clinical Estimates with Test Scores on Mental Ability and Personality Tests," *J. Clin. Psychol.*, **10**, 97-99 (1954)
179. Wolking, W. D., "Predicting Academic Achievement with the Differential Aptitude and the Primary Mental Abilities Tests," *J. Appl. Psychol.*, **39**, 115-18 (1955)
180. Yates, A. J., "The Validity of Some Psychological Tests of Brain Damage," *Psychol. Bull.*, **51**, 359-79 (1954)
181. Zimmerman, W. S., "The Influence of Item Complexity upon the Factor Composition of a Spatial Visualization Test," *Educ. Psychol. Measurement*, **14**, 106-19 (1954)
182. Zubin, J., "The Measurement of Personality," *J. Counseling Psychol.*, **1**, 159-64 (1954)
183. Zubin, J., "Failures of the Rorschach Technique," *J. Projective Techniques*, **18**, 303-15 (1954)
184. Zubin, J., "Presidential Address—Biometric Methods in Psychopathology," *Proc. Am. Psychopathol. Assoc.*, **9**, 123-43 (1954)