noted about the standardisation of these tests. The numbers used are substantial, usually exceeding 10,000 for each test, but they are drawn from selected areas and usually from ages 10-2 to 11-5. No indication is given of the method of selecting the areas, but it seems almost certain that these were the areas using the test for allocating pupils to secondary schools. Many of the pupils taking the tests were probably highly motivated to obtain high scores and gain admission to grammar schools. If the norms obtained from these pupils are used for pupils who are not so highly motivated, the resulting scores may well be depressed by several points.

A factor operating in the opposite direction is the steady improvement of scores in tests of this type that has been shown over the years,[1] so that the norms especially for the earlier tests require adjustment. The norms for those aged 11-6 to 12-0 are obtained by extrapolation. Finally, the norms are given for boys and girls taken together, although the manual points out that girls, on average, score about three points higher than boys. The recommendation in the manual is that in an allocation procedure the two sexes be considered independently of each other.

No validity coefficients are given, the reader being referred to Yates and Pidgeon's *Admission to Grammar Schools* for evidence to support the claim that the validity of verbal tests as predictors of success in the grammar school has been firmly established from empirical evidence in follow-up studies. This was without doubt a justifiable claim in the days when the grammar school was the chosen method of educating the abler children of England and Wales. The position is constantly changing, however, as the number of schools described as comprehensive increases. What function would be served by tests of this type if the system became fully comprehensive? There is evidence that in that situation tests of this kind are still satisfactory predictors of success in the examinations that must be passed before entry is gained to academic courses in universities and similar institutions. It may well be that these tests will continue to be used by administrators and teachers to enable them to advise pupils and their parents on the selection of courses of study in the secondary school.

1 PILLINER, A. E. G.; SUTHERLAND, J.; AND TAYLOR, E. G. "Zero Error in Moray House Verbal Reasoning Tests." *Brit J Ed Psychol* 30:53–62 F '60. *

For a review by Arthur B. Royse of earlier tests, see 6:511.

[399]

★**WLW Culture Fair Inventory.** Job applicants; 1969; CFI; intelligence; 1 form (8 pages); preliminary manual (8 pages); reliability and validity data based upon earlier experimental forms; preliminary norms; $15 per set of 100 tests and manual, postpaid; specimen set not available; [30–45] minutes; Barbara O. Murray (test), Lynde C. Steckle (test), and Robert W. Henderson; William, Lynde & Williams. *

ARTHUR R. JENSEN, *Professor of Educational Psychology; and Research Psychologist, Institute of Human Learning; University of California, Berkeley, California.*

The *WLW Culture Fair Inventory* (CFI) is a nontimed, nonverbal test intended to eliminate verbal and other cultural effects. The 30 items comprising the test consist of figural materials involving logical-spatial relationships. There are five parts: (*a*) selection of the one figure in a set of five that is most different, (*b*) block counting, (*c*) selection of the fifth figure in a series, (*d*) paper form board-type items, (*e*) selection of a figure which completes a pattern (matrices). Each of these parts is preceded by printed instructions and two examples with the correct answers given.

Judging from inspection of the items and this reviewer's experience with a wide variety of psychometric tests, it is inferred that the CFI measures reasoning and spatial abilities, much the same abilities as are assessed by Cattell's *Culture Fair Intelligence Tests,* Raven's *Progressive Matrices,* and Domino's *D48 Test,* although the CFI appears to be more loaded on a spatial factor.

The test manual claims that the test should "be suitable for minority groups or 'hard-core' individuals in hiring situations." The test, however, appears to be too difficult for this population. The 50th percentile for college students is only 18 items correct (out of 30), and the median for Negro clerical workers is only 10 items correct; the 99th percentile for this group is only 20 items correct. It is this reviewer's prediction that culturally disadvantaged minority persons will average lower on this test, in relation to middle class whites, than on most other standard tests, especially verbal tests.

The test has at least three advantages: (*a*) it is "nonverbal" and eliminates or minimizes the influence of reading ability and formal education; (*b*) it presents a variety of figural subtests, so that its total score should not con-

tain a factor specific to a particular form of figural test material such as Matrices; (*c*) it is untimed and is thus a power test, but this also may introduce personality factors into the score variance. (The reviewer found a correlation of −.45 between time taken on the *Progressive Matrices* and the extraversion scale of the *Eysenck Personality Inventory* when subjects had no time limit. The more extraverted subjects tend to get through faster or to give up sooner when the items increase in difficulty.) Raven's PM shares advantages *a* and *c*; Cattell's CFT shares advantages *a* and *b*.

Several criticisms can be made of the CFI: (*a*) Some of the items depend as much upon perceptual or visual acuity as much as upon reasoning ability. (*b*) The figural materials are not as clearly printed as those in the Cattell and Raven tests. (*c*) The test is too difficult for the general population, being more suited to college level persons. (*d*) There are no really easy items at the beginning of each subtest to permit the subject to catch on readily to what is required and to gain practice before he is confronted with the more difficult problems. The Cattell and Raven tests are superior in this respect.

Finally, the test manual is totally inadequate and makes it obvious that the test has been published and marketed much too prematurely. The manual contains virtually no helpful information. The "norms" of the CFI are based on 50 male and 50 female college students, and 24 Negro and 46 women applicants for clerical jobs in an insurance company! The intercorrelations among the five subtests are remarkably low, ranging from .29 to .55. No data are given concerning the test's reliability. Correlations with other intelligence tests are so low as to make one wonder what the CFI actually measures. Aside from these correlations suggesting very low concurrent validity, the manual gives no other evidence of the test's validity or usefulness.

In conclusion, there is no good reason to recommend this test in its present inadequate stage of development. Testers who are seeking a good nonverbal "culture fair" test of general intelligence are urged to consider Cattell's *Culture Fair Intelligence Tests* or Raven's *Progressive Matrices* (these tests come in different forms for various ages and populations). The construction of these alternative tests, the adequacy of their manuals, their normative data, and the research behind them are totally unmatched by the *WLW Culture Fair Inventory.*

JAMES E. KENNEDY, *Professor of Psychology, The University of Wisconsin, Madison, Wisconsin.*

This 30-item test, consisting of five varieties of pictorial problems, requires no writing and was designed for group administration with no time limit. The test manual does not say what the test purports to measure. Since most of the items were adapted from or modeled after items from standard intelligence tests, one can infer that it was probably intended to measure one or more aspects of nonverbal reasoning or intelligence.

The test was developed to meet the expressed need of industrial clients for "a culture fair test that would largely discount the influence of environment and education." Toward this end, the test is untimed and limited to nonverbal items which "most people have not encountered in their daily lives." It is said to be suitable for use with "minority groups or 'hard-core' individuals in hiring situations."

The test manual offers only a vague and clumsy description of how the test was developed. Using a sample of students in a small women's liberal arts college, some type of item analysis was performed on the nonverbal section of the WLW Mental Alertness Inventory, the *Culture Fair Intelligence Test,* and the *Chicago Non-Verbal Examination* as a basis for the selection of 70 items. In some fashion or other these 70 items were reduced to the 30 items comprising the current form of the test. Reference is made to adopting items from other tests as well, but there is no explanation. (The manual does not explain why the authors thought it advisable to use an item-analysis sample of college students to develop a test for "minority groups or 'hard-core' individuals.")

Norms are available on 50 male and 50 female college students and on 24 black and 46 white female applicants for entry clerical jobs at an insurance company. No further description of these groups, such as educational level, age, or socioeconomic status, is provided.

The manual also presents a hodgepodge of correlations between part scores on the *WLW Culture Fair Inventory* and part or whole scores on several other tests. Since the WLW scores used were based on the preliminary 70-item form of the test, these correlations have little

relevance for the test being reviewed. No reliability estimates of the test are reported.

The manual reports a median score of 10 for the black applicants and 13 for the white applicants on this test and a median score of 20 for the blacks and 24 for the whites on the *Wonderlic Personnel Test*. While information of this kind is meaningless without a discussion of the comparability of these groups on other dimensions, it would not appear that this purportedly culture fair test operates differently from the Wonderlic, which makes no claims about culture fairness.

In summary, this test is not based on any stated theoretical grounds, the operations followed in its development are obscure, its reliability has not been estimated, and its validity has not been demonstrated nor even considered in the manual. The assumption that elimination of the time and verbal factor would "largely discount the influence of the environment and education" was at least naive, and, in my opinion, absurd. I know of no test that would serve the purpose for which this test was designed.

[400]

***Wesman Personnel Classification Test.** Grades 8–16 and adults; 1946–65; WPCT; title on Forms A and B is *Personnel Classification Test;* 3 scores: verbal, numerical, total; Forms A ('46, 4 pages), B ('47, 4 pages), C ('64, 4 pages); revised manual ('65, 27 pages); $3 per 25 tests; 50¢ per specimen set; postage extra; 28(35) minutes; Alexander G. Wesman; Psychological Corporation. *

REFERENCES

1–3. See 4:331.
4–11. See 5:399.
12. ASH, PHILIP. "Validity Information Exchange, No. 13-05:D.O.T. Code 1-86.12, Salesmen, Typewriters." *Personnel Psychol* 13:454 w '60. *
13. ASH, PHILIP. "Validity Information Exchange, No. 13-06:D.O.T. Code 5-83.127, Typewriter Serviceman." *Personnel Psychol* 13:455 w '60. *
14. DUNNETTE, MARVIN D., AND KIRCHNER, WAYNE K. "Psychological Test Differences Between Industrial Salesmen and Retail Salesmen." *J Appl Psychol* 44:121–5 Ap '60. * (*PA* 35:4029)
15. VINCENT, NORMAN L., AND DUGAN, ROBERT D. "Validity Information Exchange, No. 15-03:D.O.T. Code 1-57.10, Salesman, Insurance." *Personnel Psychol* 15:223–5 su '62. *
16. DUNNETTE, MARVIN D.; WERNIMONT, PAUL; AND ABRAHAMS, NORMAN. "Further Research on Vocational Interest Differences Among Several Types of Engineers." *Personnel & Guid J* 42:484–93 Ja '64. * (*PA* 39:6040)
17. WELSCH, LAWRENCE A. *The Supervisor's Employee Appraisal Heuristic: The Contribution of Selected Measures of Employee Aptitude, Intelligence and Personality.* Doctor's thesis, University of Pittsburgh (Pittsburgh, Pa.), 1967. (*DA* 28:4321A)
18. PARRY, MARY ELLEN. "Ability of Psychologists to Estimate Validities of Personnel Tests." *Personnel Psychol* 21: 139–47 su '68. * (*PA* 42:14727)

ARTHUR C. MACKINNEY, *Dean of Graduate Studies, Wright State University, Dayton, Ohio.*

Everyone has come to expect high quality from both Wesman and the Psychological Corporation, and this test maintains the tradition. Overall, it appears to this reviewer to be an excellent example of a short, general-population-level, general intelligence test, composed of verbal and numerical subtests.

The verbal subtest is composed of 40 two-part analogy items arranged in increasing difficulty. The numerical subtest is composed of 20 items, mainly arithmetic, progressing from simple addition through more complex manipulations of fractions, square roots, and the like. Scores from both subtests are summed for a total score.

The manual is clear and well written, simple without being insulting. It is cautious and scholarly to an appropriate degree without compromising the obvious need for utility. Administration instructions are lucidly stated. Twenty-nine sets of norms are presented, encompassing many different occupational groups (ranging from "production workers" to "executive trainee" and "high level positions"), as well as several student groups (ranging from tenth grade students to college sophomores). Reliabilities (including the sub-score reliabilities), equivalence of forms, correlation between subtests and the influence of speed (negligible) are all within the acceptable range. Validity data of three general types are presented. First, correlations of this test with other tests give some good indication of the conceptual content being tapped. Second, the progression of mean scores across the occupational hierarchy is clearly demonstrated. Third, data from 12 concurrent validity studies show correlations primarily under .50. In general, these seem to follow the typical pattern of higher correlation for higher level jobs and for the more intellectual kinds of criteria. The user can be confident that the validity of this instrument will be no worse than that of similar tests, and probably better. In general, the use of this instrument seems indicated in those instances for which it was designed.

J Counsel Psychol 12:435–6 w '65. Jack C. Merwin. * Using two groups, one a high scoring group and one a lower scoring group, the proportion completing nine-tenths the items of each subtest in the stated time limits was calculated. * The proportion of the low scoring group completing nine-tenths of the numerical test, however, was only 43 per cent. The author interprets this as, "presumably a matter of running out of ability rather than out of time."