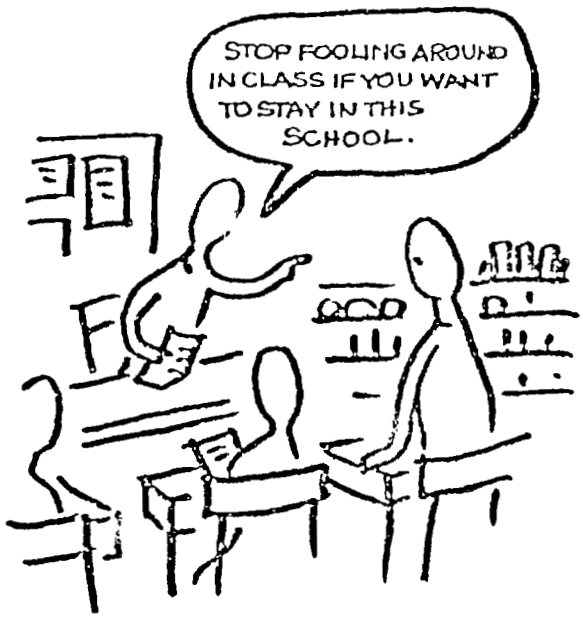


"You don't scare me."

How many students do you think would give this answer to a teacher?

- ☐ 2 out of 10 students
- ☐ 3 out of 10 students
- ☐ 4 out of 10 students
- ☐ 5 out of 10 students
- ☐ 6 out of 10 students

Help sick people in hospitals keep busy by showing them how to make things.

- ☐ Pretty bad - I couldn't take doin
- ☐ Not too good - I might do this bu wouldn't like it.
- ☐ This is OK - I wouldn't mind this too much.
- ☐ This is great - just the kind of I would like to do.

Journal of Educational Psychology
1970, Vol. 61, No. 3, 229-240

ASSESSMENT OF DISADVANTAGED ADOLESCENTS: A DIFFERENT APPROACH TO RESEARCH AND EVALUATION MEASURES[1]

NORMAN E. FREEBERG[2]
Educational Testing Service, Princeton, New Jersey

The suitability of available formal tests for individuals from culturally deprived and poverty-level backgrounds has long been open to criticism. The present study presents an attempt to overcome a number of the defects customarily attributed to conventional measures by the construction and application of a test battery tailored specifically in content, format, and administration to disadvantaged adolescent groups. The measures were administered to 256 male and female high school dropouts enrolled in a federally funded, youth-work training program. Results based upon item characteristics, levels of internal consistency, validity (against rating criteria), the interpretability of the factor patterns, and various research findings suggest that the measures—as well as the techniques utilized in their design—warrant further development and research application with disadvantaged adolescents.

The advent of large-scale, federally funded, work-training programs to assist disadvantaged young adults in making a suitable vocational and social adjustment has brought with it a need for objective measures applicable to individuals from poverty-level backgrounds. Available formal tests of intellectual skills and attitudes that might be used for individual classification or evaluation purposes can only be recommended with caution, for reasons that have been well summarized (Karp & Sigel, 1965; Society for the Psychological Study of Social Issues, 1964). Adverse effects on test scores have been attributed to such influences as low verbal reading levels, poor test-taking motivation, failure to comprehend middle-class cultural content or language usage, and negative attitudes toward academic trappings and toward an activity which has little discernible "payoff" for the respondent. Additional influences claimed include a lack of test-taking skills, poorly designed test formats or instructions, and adverse examinee attitudes toward the examiner, especially when he is of a different race.

Agreement that complex problems exist is not necessarily reflected in unanimity of the remedies proposed. Attempts to develop culture-fair tests as a solution have long been open to question in terms of success at achieving culturally "neutral" content or suitable predictive validity (Anastasi, 1950, 1965; Eels, Davis, Havighurst, Herrick, & Tyler, 1951; Wesman, 1968). In educational and employment settings, where test "bias" against disadvantaged minority group applicants has been claimed, the evidence and the proposed solutions have generally focused upon differential predictability. Relevant data derived from conventional tests administered to various ethnic or cultural groups have been presented (Cleary & Hilton, 1968; Kirkpatrick, Ewen, Barrett, & Katzell, 1967; Lopez, 1966; Potthoff, 1966); but this approach unfortunately is based upon use of *existing* test materials and fails to correct for the flaws in test design attributed to such measures. A third, but less clearly defined, solution has merely stressed the need for "caution" and "flexibility" when conventional test materials are applied to persons from disadvantaged backgrounds (Society for the Psychological Study of Social Issues, 1964).

An alternative and more effective approach may be sought in designing measures that specifically attempt to overcome

[1] This research was conducted under contract with the Bureau of Work Training programs, United States Department of Labor (Contract No. 66-00-09).

[2] Requests for reprints should be sent to Norman E. Freeberg, Developmental Research Division, Educational Testing Service, Princeton, New Jersey 08540.

generally agreed upon deficiencies of conventional test materials and techniques, when applied to the culturally disadvantaged. Although the suggestion is hardly novel (Gordon, 1963; Karp & Sigel, 1965; Lennon, 1964), little has been done to develop and apply such measures, to demonstrate their acceptability in terms of psychometric properties, or to determine their value as research tools. The present study is an initial attempt to develop a set of measures for use in evaluation of youth-work programs that are based on these objectives. The rationale supporting this approach is that much of the material in conventional formal measures, as well as the mode and format of presentation, is designed primarily for individuals from a middle-class cultural background and as a result the material fails to communicate much of its intended meaning (verbally or experientially) to those who are culturally different. Failure of the respondent to grasp intended concepts can result in questionable response patterns. A more appropriate (valid and reliable) behavior sample is likely to be elicited when more culturally relevant stimulus materials are presented.

METHOD

Description of the Measures

To provide some understanding of the techniques utilized in implementing the above rationale, the general characteristics of a new battery of measures are described below.

Format. The items are printed in small booklets (approximately 5 × 8 inches) with the items and response alternatives always appearing on a single or adjoining page. This format is intended to minimize the confusion claimed when students transfer responses from a page of items to a separate answer sheet (Clark, 1968). Seven of 13 test booklets used pictorial information accompanied by supplementary verbal information. The remaining six booklets are entirely verbal. Separate forms for males and females are required for those measures that contain pictorial information.

Administration. The measures are intended for administration to small groups of approximately 12 individuals. Directions and all item stems and response choices are *orally* presented by the examiner. There are no overall time restrictions, since the examiner paces the presentation by reading the test items.

Verbal content. The difficulty of the verbal material was maintained at approximately a fifth-grade reading level, based on a check against the Thorndike-Lorge Word List (Thorndike & Lorge, 1944). Slang expressions, generally suitable for adolescents, are used where they are considered helpful in clarifying item content. Similarly, the subject matter was chosen so as to be relevant to, and comprehended by, adolescents from a disadvantaged or low-socioeconomic-status (SEC) background.

Pictorial content. For each item utilizing pictorial information, the scene presented portrays a youthful adolescent figure in some situation relevant to the particular variable being measured (e.g., responding to an authority figure). Pictorial information is intended to clarify verbal content and to enhance test-taking motivation with more "interesting" test materials (Geist, 1964). Since the measures are intended to be applicable to a number of racial or ethnic groups and the respondent is expected to identify with the youthful figure in a pictorial scene, any features that depict race were "neutralized" by using simple line drawings in a cartoon style similar to those used in the Rosenzweig Picture Frustration Test (Rosenzweig, 1948).

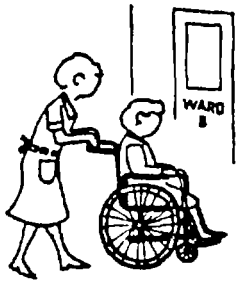
Test length. The "test-taking attention span" of the intended respondents can be expected to be relatively short. Tests were made only long enough to provide usable levels of reliability. The number of items in each test booklet ranges from a minimum of 10 to a maximum of 30.

A brief description of specific test content is also needed in order to understand the results obtained. The tests are intended to measure aspects of trainee behavior pertinent to generally agreed-upon goals for a youth work-training program. These are:

Job Knowledge. Various jobs are pictured along with their job titles and accompanied by multiple-choice items covering six knowledge categories: (*a*) education required, (*b*) where work is performed, (*c*) full-time starting salary, (*d*) the specific, primary task performed, (*e*) working hours required, and (*f*) tools used in perfoming the job (with response choices consisting of pictured tools).

Vocational Aspirations. Each item pictures a different occupation, accompanied by a job title and a one-sentence description of a major task required for that job (see Figure 1). Jobs were chosen so as to be both relatively familiar and scorable on a continuum of occupational status or prestige (Duncan, Hatt, & North, 1961). Two items at each of eight status levels from low (Status Level 1—Laborer, Porter) to high (Status Level 8—Scientist, Physician) make up the 16 items of the measure. The respondent indicates, on a 4-point scale, the degree to which he would *like* to enter each occupation. Scoring is based upon consistency in judging "desirability" of job pairs of the same status levels.

Vocational Plans. This test presents the same



HOSPITAL ATTENDANT

Helps take care of sick people in a hospital.

- ☐ A great job; the kind of thing I would like.
- ☐ It's pretty good; I wouldn't mind doing this.
- ☐ Not the greatest -I'd do this if there weren't anything better.
- ☐ This is a rotten kind of job. I wouldn't like it at all.

FIG. 1. An example of an item from Vocational Aspiration Measure. (Reprinted with permission of the Educational Testing Service, copyrighted 1967.)

material on occupations as the Vocational Aspiration measure and is scored in the same way, except that response choices reflect the degree to which the respondent would actually *expect* (i.e., plan) to enter the particular occupation. A difference score was also obtained by summing the differences between Vocational Aspiration and Vocational Plans scale values assigned to the same occupations.

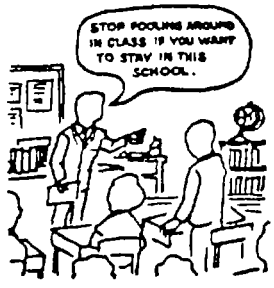
Interest in Vocational Tasks. There are seven interest clusters of four items each. The respondent indicates the degree of interest in specific job tasks that are both pictured and described verbally. The clusters are based upon factor analytic findings used in the development of the California Occupational Preference Survey (Knapp, Grant, & Demos, 1966). Each cluster is scored separately.

Attitude toward Authority. The test presents scenes showing an adolescent being addressed by an authority figure (e.g., judge, supervisor, teacher, policeman, father, mother). A response made by an adolescent which is either clearly pro-authority or anti-authority in nature is presented below the picture (Figure 2). The 5-point response scale is intended to reflect the respondent's "direction of perception" (Bernberg, 1954) in relation to the pro-authoritarian or anti-authoritarian response presented.

Self-Esteem. This test presents 16 pictorial scenes in which an adolescent is portrayed in various academic, social, and employment settings. The respondent is to imagine himself as the youthful person in the scene and to indicate on a 3-point response scale what the outcome might be. The item responses are intended to reflect a form of self-esteem often conceived of as competence in interpersonal or social action and the ability to affect one's social environment.

Deferred Gratification. This is an all verbal 20-item measure intended to span several dimensions that have come to be known as the deferred gratification pattern (Schneider & Lysgaard, 1953; Straus, 1962). The pattern deals with items that define a willingness to delay present reward for future gain. Four subscales of four items each are used to derive the total scores and are designated as Hasty Aggression, Freedom of Spending, Affiliation, and Delay of Reward.

Job Seeking and Job Holding Skills. The test consists of two separate item types combined into a single booklet but analyzed separately. The first set of 17 items is in a 4-choice format and deals with questions about where to look for a job and how to interpret the information in a newspaper want ad and in portions of a job application blank. The second set (11 items) pictures



"You don't scare me."

How many students do you think would give this answer to a teacher?

- ☐ 2 out of 10 students
- ☐ 3 out of 10 students
- ☐ 4 out of 10 students
- ☐ 5 out of 10 students
- ☐ 6 out of 10 students

FIG. 2. An example of an item from the Attitude toward Authority Measure. (Reprinted with permission of the Educational Testing Service, copyrighted 1967.)

situations that might be encountered on the job. These cover such customary employer concerns as punctuality, appropriate dress, response to a supervisor's requests, etc.

Motivation for Vocational Achievement (I and II). There are 17 verbal items in a 5-point Likert-scale response format. The first 10-item stems (Motivation for Vocational Achievement I) are statements about obtaining and maintaining employment (e.g., "It pays to make something of yourself by working"). The 7 remaining items (Motivation for Vocational Achievement II) are conditional statements about willingness to maintain a job (Would you stay with it if: "The work was dirty?", "You had to work long hours?", etc.).

Practical Reasoning. There are three separate measures that attempt to deal with job-related reasoning (i.e., direction-following) tasks. The three 10-item booklets each contain 4-choice items and are designated as follows:

1. *Map Reading* requires the respondent in an assumed job to utilize a map for traveling to various customers. Items are concerned with number of blocks to be covered, direction traveled, etc., in reaching the destination.
2. *Zip Coding* presents items dealing with the respondent's skill at "sorting" mail in an assumed postal job, using zip-code type of information presented on a chart included in the test booklet.
3. *File Card Sorting.* The respondent is asked to sort out individual applicants for jobs, based upon the characteristics presented on a set of cards. The correct responses are dependent upon various job requirements stipulated in item stems (e.g., a woman in a particular age range with given educational level, given skills, etc.).

Enrollee Rating Scale. This form consists of a 15-item evaluation, by the trainee, of various aspects of the Neighborhood Youth Corps Program in which he is enrolled.

Sample

Data were obtained from administration of the test battery to samples of 123 males and 133 females enrolled at 11 Neighborhood Youth Corps project centers located in urban areas of the northeast United States.[3] Approximately 85% of the male and 90% of the female enrollees in the sample were Negro; the remainder were white or of Puerto Rican ethnic background. All of the subjects were school dropouts between 16 and 21 years of age who had been enrolled in what is designated as the Neighborhood Youth Corps "Out-of-School Program" for at least 1 month.

[3] Neighborhood Youth Corps is a federally funded, national youth-work training program for disadvantaged adolescents administered by the United States Department of Labor and the United States Office of Economic Opportunity (United States Department of Labor, 1965).

Criterion Measures

Two rating scales of 11 items each served as the criterion measures for the study. One scale was designed for use by guidance counselors and the other for use by job-site supervisors who are the program personnel having closest daily contact with the trainee. Each trainee is assigned to a counselor at the project center, who determines job placements and provides overall vocational guidance services. A work supervisor is responsible for the trainee's performance on the assigned job. Jobs range from a helper in the laundry room of a hospital to an assistant in the electronics repair shop of a military base. The first 10 items used in each scale were based upon aspects of trainee behavior derived from interviews with counselors (e.g., "He is coherent in expressing himself") and work supervisors (e.g., "He gets to work on time"). The eleventh item of each scale represented an overall evaluation of the trainee's performance and adjustment in the work-training program. Each rater generally completed evaluations of several trainees for whom he was responsible.

Data Analysis

The analyses used are intended to define psychometric characteristics of the individual tests and to suggest ways of improving and using them in future research applications. The primary analyses consist of computing (*a*) item analyses (*p* values and item-total test correlations), (*b*) internal consistency estimates (reliability), (*c*) validity coefficients, (*d*) intercorrelations, and (*e*) the factor patterns, with criteria added to these factors by extension. Sample sizes varied slightly for the different measures, as indicated in the tables, thus requiring missing data analyses.

Results

Item Analyses

The analyses of 225 items that make up the battery provided the basis for eliminating those considered unsuitable.[4] Elimination was based upon (*a*) particularly discrepant mean values (i.e., *p* values greater than .90 or extremely high or low mean values for scaled items), (*b*) minimal variance, and/or (*c*) item-total test correlations tending to be zero or negative. Relatively few items were found unacceptable. A total of only five items

[4] For these item data, order NAPS Document 00860 from ASIS National Auxiliary Publication Service, % CCM Information Sciences, Inc., 909 3rd Avenue, New York, New York 10022, remitting $1.00 for microfiche or $3.00 for photocopies.

appearing in two measures (Job Knowledge and Self-Esteem) were dropped from use with the male sample. Eight items appearing in four measures used with the female sample were dropped, plus one complete measure—the Practical Reasoning File Card Sorting subtest—which was far too easy for the females.

In general, there was little evidence of the excessive variance restriction or homogeneity often attributed to item results when conventional measures are used with disadvantaged groups. As will be seen, the item variability was sufficient to produce measures which had a number of logically interpretable correlations of substantial magnitude with other variables.

Reliability

The internal consistency estimates, shown in Table 1, were based upon one of the three methods of analysis indicated in the footnotes to the table.

For the most part, internal consistency is substantial despite the relatively few items in each measure. The primary exception is found for the overall reliability of the Deferred Gratification measure (r's = .37 and .29). Internal consistency estimates of each of the four subscales that make up the total Deferred Gratification measure (Table 1) indicate that the possible reason lies in the poor reliability of the subscale designated as Delay of Reinforcement, which obviously does not form a coherent scale. The reliabilities for Vocational Plans and Aspirations are based upon split-half correlations using matching job-status pairs in each half. Of interest are the reasonably high levels of consistency shown in selecting jobs on a status basis, especially for the males.

TABLE 1
RELIABILITY ESTIMATES FOR ENROLLEE MEASURES

Enrollee measure	Males		Females	
	No. of items	(r_{tt})	No. of items	(r_{tt})
Job Knowledge[a]	30	.75	27	.72
Vocational Plans[b]	16	.81	16	.57
Vocational Aspirations[b]	16	.78	16	.67
Vocational Interest (total)	28		28	
Clerical[c]	4	.62	4	.80
Service[c]	4	.69	4	.65
Technical[c]	4	.39	4	.71
Outdoor[c]	4	.65	4	.78
Science[c]	4	.60	4	.69
Business[c]	4	.56	4	.64
Aesthetic[c]	4	.67	4	.45
Attitude toward Authority[b]	13	.64	12	.87
Self-Esteem[b]	16	.50	16	.60
Deferred Gratification[b] (total)	16	.37	16	.29
Freedom of Spending[c]	4	.74	4	.69
Hasty Aggression[c]	4	.64	4	.56
Affiliation[c]	4	.61	4	.61
Delay of Reinforcement[c]	4	.07	4	.06
Job Seeking Skills[a]	17	.65	17	.66
Job Holding Skills[b]	11	.74	11	.72
Motivation for Vocational Achievement				
MVA-I[b]	10	.81	10	.58
MVA-II[b]	7	.81	7	.64
Practical Reasoning (total)	30		30	
Map Reading[a]	10	.60	10	.47
Zip Coding[a]	10	.75	10	.66
File Cards[a]	10	.71	10	—[d]
Enrollee Rating Scale[b]	15	.96	15	.90

[a] Kuder-Richardson (21) estimates.
[b] Split half; stepped up by Spearman-Brown formula.
[c] Average interitem r stepped up by Spearman-Brown formula.
[d] Dropped from use with female sample.

Validation and Associated Research Findings

The validity coefficients presented in Table 2 are low to moderate in size.[5] For male enrollees, consistently significant validities are found for the Practical Reasoning subtests, with the highest validities for work supervisor ratings. Male enrollees who have the ability to follow directions in carrying out a variety of

[5] It is of interest that the levels of validity, in the initial use of these measures, compare favorably with the approximate median validity of .22 reported for the General Aptitude Test Battery—against criterion ratings of job proficiency—after 20 years of use by the USES (Bemis, 1968). That program is probably the only one involving long-term systematic test research that has included samples of disadvantaged youth similar to those used in the present study.

TABLE 2

CORRELATIONS OF COUNSELOR AND WORK SUPERVISOR CRITERION RATINGS WITH ENROLLEE MEASURES

Enrollee measure	Males		Females	
	Counselor ($n \cong 112$)	Work supervisor ($n \cong 102$)	Counselor ($n \cong 129$)	Work supervisor ($n \cong 128$)
Job Knowledge	.22*	.24*	.19*	.03
Attitude toward Authority	.12	.22*	.00	.06
Self-Esteem	.04	.01	.15	−.21*
Deferred Gratification	.27**	.17	.18	−.03
Job Seeking Skill	.32**	.35**	.26**	.22*
Job Holding Skill	.22*	.19	.26**	.17
Motivation for Vocational Achievement I	.26**	.21*	.21*	.13
Motivation for Vocational Achievement II	.02	.03	.10	.13
Practical Reasoning				
Map Reading	.24*	.28**	.06	.00
Zip Code	.20*	.37**	.28**	.08
File Card	.23*	.31**	—	—
Enrollee Rating Scale	.22*	.21*	.20*	.15
Vocational Aspiration	−.05	−.03	−.03	−.13
Vocational Plans	−.04	−.06	−.11	−.13
Vocational Aspiration—Vocational Plans (Difference)	−.07	−.16	.00	−.07
Vocational Interest				
Clerical	.17	.16	.10	.03
Service	.04	−.01	.25**	−.01
Technical	−.11	−.02	.00	−.01
Outdoor	.04	.01	.08	.01
Science	.09	.09	.16	.00
Business	.07	.14	−.04	−.09
Aesthetic	.04	−.12	.07	−.19

* r significantly greater than zero at .05 level.
** r significantly greater than zero at .01 level.

job-related reasoning tasks tend to be seen as superior by the work supervisors. The only other measures that reflect a distinctly cognitive component are Job Knowledge and Job Seeking Skills, which also achieve consistently significant validity against both criteria. For females, the measures having a cognitive or intellectual component tend to attain significant validity primarily with the counselor ratings. Among the attitudinal type of measures, those that are generally superior in validity for both subject groups deal with attitudes toward working and toward the training program (Motivation for Vocational Achievement I and the Enrollee Rating Scale).

An appreciation of limitations in the concurrent criterion measures helps to provide an understanding of the validity levels achieved and ways in which these might be improved. First, despite fairly specific behavioral items that make up the scales, a considerable halo effect resulted. This was evident in such unusually high intrarater reliability coefficients (split-half) as .94 for counselor ratings of males and .98 for work supervisor ratings of females. In addition, the final item of each scale (single summary evaluation) produced item-total scale correlations in the .80s. Excessive leniency in rating, as evidenced by high item and total scale means, was shown particularly in ratings of female trainees by work supervisors. The effect produced considerable restriction on item variance and can help explain why there were few significant coefficients obtained with work supervisor criterion ratings of females. Further evidence that the quality of the ratings might be improved was found in a rather low level of interrater agreement ($r = .36$ for both male and female samples). Counselors' and work supervisors' "successful" or "unsuccessful" enrollees obviously do not perceive in the same way. Determination of the accuracy of these evaluators against longer range trainee performance criteria would be in order for the eventual development of more effective criterion scales.

Table 2 shows a uniform lack of significant validity for those vocational measures dealing with interests, plans, and aspirations. Some attenuation in validity for the vocational interest clusters is probably attributable to the small number of items. However, there is also the question of whether these groupings of occupational tasks, derived from the California Occupational Preference Survey (Knapp, et al. 1966) and based upon a predominantly white middle-class sample, are equally applicable when pictorial information is added and the items are used with a sample of disadvantaged youngsters. Factor analyses of the present VI-item data, using principal components solutions and Varimax rotations (Kaiser, 1958), resulted in reclustering of the interest items in several important ways when compared with the California Occupational Preference Survey. In brief, the analyses indicated that only the Clerical cluster remains intact. The Business, Aesthetic, and Science occupational clusters were found to lose their coherence and differentiate into factors designated as High

TABLE 3
VOCATIONAL ASPIRATIONS AND PLANS MEAN SCORES BY STATUS LEVELS

Status Level	Aspirations		Plans	
	Male[a]	Female[b]	Male[a]	Female[b]
1	1.47	1.49	1.64	1.33
2	2.18	2.09	2.07	1.74
3	2.28	2.37	2.04	1.75
4	2.62	2.57	2.27	2.27
5	2.94	2.92	2.66	2.54
6	2.98	3.31	2.40	2.65
7	2.66	3.26	1.98	2.31
8	3.22	2.74	2.01	1.65

[a] $n = 122$.
[b] $n = 133$.

Level Technical and Business, Lower Level Technical Skills, and a Personal Service factor (that involves offering services to people on a face-to-face basis in any sort of setting). In addition, there is a distinctly Low Level Occupations cluster defined by items covering menial, outdoor tasks and lower level technical jobs (e.g., "running a pressing machine"). The perception of job tasks would thus appear to be more highly influenced by status considerations in a sample composed exclusively of disadvantaged adolescents than is the case for a wider cross-section of young adults.

The mean scores at each of the eight job-status levels provide direct evidence of status considerations in the group's vocational perceptions (Table 3).[6] The values are found to be reasonably close in their conformity to a conventional status continuum (especially among males). Also evident is a fairly uniform drop in mean job choice scores when going from aspirations to plans, which may be interpretable as a "reality" shift. The results are consonant with the findings of Stephenson (1955) and Empey (1956) in confirming the importance of dealing with "preferred" and "anticipated" choices explicitly when studying vocational desirability. An additional feature in dealing with scores on job-choice measures is evident when the mean item responses for Vocational Plans and Vocational Aspirations are grouped into categories of low status (Prestige Levels 1–4) and high status (Levels 5–8), and the difference between mean differences at these two levels is tested (Table 4). The highly significant difference that results makes it strikingly clear that, at least for a disadvantaged adolescent group, the joint use of scores from any plans and aspirations measures must take into account the status levels of the vocations presented. In effect, this result indicates that the youngster does not aspire to be a low-status porter or laborer; nor would he plan to enter such occupations. On the other hand, he would aspire to be a high-status physician or scientist, but is less likely to consider it probable (i.e., plan) that he would enter those occupations. Differential weighting of scores, or construction of separate scales based on high and low status occupational groupings, are suggested by such a result.

One other vocational orientation measure that lends itself to closer examination is Job Knowledge. The mean proportion of correct responses for each knowledge category is shown in Table 5. (Items dealing with tools used for the job had been elimi-

TABLE 4
TEST OF DIFFERENCE BETWEEN MEAN DIFFERENCES FOR PLANS AND ASPIRATIONS AT HIGH AND LOW STATUS LEVELS

Status	Vocational Aspirations M	Vocational Plans M	M_{diff}
Male[a]			
High	2.95	2.26	.69[c]
Low	2.14	2.02	.12[c]
Female[b]			
High	3.06	2.29	.77[d]
Low	2.13	1.77	.36[d]

[a] $N = 122$.
[b] $N = 133$.
[c] CR mean difference = 8.22, $p < .001$.
[d] CR mean difference = 6.97, $p < .001$.

[6] The abrupt shift downward for the highest status level jobs (Level 8) among females is considered somewhat of an artifact in that the highest, or professional, level jobs tend to be those which are perceived as exclusively occupied by males and not realistically within the province of females (e.g., physician, lawyer, judge, scientist, engineer, etc.).

nated from any analyses because of virtually perfect scores achieved.) The ordering of difficulty levels should be interpreted cautiously, however. "Place" where work is performed produces scores with particularly low mean difficulty levels, but this may result from difficulty in devising suitable distractors for the question, "Where would you work on this job?" when using relatively conventional occupations. The problem is similar, but not as acute, for the "Nature of Work" category. Some limited evidence that the three remaining categories may in fact reflect job knowledge levels more accurately, as opposed to greater ease in item construction, can be found in questionnaire and anecdotal materials that report considerable naiveté among minority group adolescents regarding training requirements and economic aspects of employment (Amos, 1964; Wispé, 1965).

Intercorrelations and Factor Patterns

A more coherent picture of the behavioral dimensions tapped by the measures of the battery can be obtained by factoring the intercorrelation matrices for the male and female samples (Table 6).

The factors were extracted from each matrix by a principal components method and those with latent roots greater than 1.00 were rotated using the Varimax technique (Kaiser, 1958). The two criterion rating scales were added to the rotated matrices by factor extension. Table 7 provides the factor designations and loadings for the samples. Loadings greater than .25 are arbitrarily utilized in the interpretation of the factors. Factor I is readily identifiable as an Intellectual Skills factor, since the highest loadings appear for those measures that require some form of information or reasoning skill. Attitude measures also achieve moderate loadings on the intellectual dimension. The negative loadings for the Outdoor cluster of the VI measure fit the interpretation of the factor, since items in the cluster depict fairly menial job tasks (e.g., farming, gardening) that tend to be unacceptable to those who score high on intellectual skills. The Intellectual Skills dimension is the only one for males on which criterion scales achieve extension loadings of interpretable magnitude; with work supervisor judgments more heavily weighted. For females, on the other hand, only Factor II, Positive Attitudes, loads on either criterion variable.

TABLE 5

PROPORTION OF CORRECT RESPONSES FOR JOB KNOWLEDGE CATEGORIES

Knowledge category	Males	Females
Place performed	.90	.95
Nature of work	.85	.88
Hours worked	.76	.78
Education required	.59	.78
Starting salary	.51	.60

Factor III, designated General Vocational Interest, results from the relatively high intercorrelations among the scores of the VI measure. For males, the factor represents a global acceptance of job tasks, with specific rejection of the low-status tasks of the Outdoor cluster. For females, it is the Clerical tasks that are unique and apparently of high status, since clerical tasks have the highest loading on the factor designated as High Level Occupational Interest (Factor V). The factor also characterizes a female who is consistent in her status perception of jobs and is highly motivated to work.

The remaining factor (Factor IV) is clearly attitudinal for each sample. Differences in the specific measures on which the highest loadings appear, however, suggest its designation as a Social Conformity dimension for females—that is, conventional attitudes toward authority and social values; while the pattern defines a Negative Self-Image for males—that is, favoring outdoor menial tasks, tendency toward a lowered self-esteem, and minimal difference between job plans and aspirations.

CONCLUSIONS

The evidence presented points to reasonable—and usable—patterns of response among disadvantaged adolescents when instruments are "tailored" to compensate for deficiencies frequently attributed to

TABLE 6

TOTAL SCORE CORRELATION MATRIX ENROLLEE MEASURES

Measure	1	2	3	4	5	6	7	8	Practical Reasoning			12	Vocational Interest							20	21	22
									9	10	11		13	14	15	16	17	18	19			
Job Knowledge (1)	—	13	19	28	58	49	46	15	45	45	46	36	13	−15	03	−26	19	12	−10	−06	−05	06
Attitude toward Authority (2)	11	—	22	30	37	40	16	−12	31	20	25	24	−09	−09	10	−04	04	10	01	−08	−08	06
Self-Esteem (3)	10	12	—	33	17	40	22	03	03	10	10	31	01	−01	11	−16	04	04	−10	−12	−12	25
Deferred Gratification (4)	35	36	07	—	22	29	38	16	31	30	31	24	01	−01	09	−02	13	13	01	−10	−39	13
Job Seeking Skill .5)	58	13	13	33	—	38	39	−00	55	60	65	29	20	−05	07	−33	28	32	−01	−25	−05	−06
Job Holding Skill (6)	29	30	40	33	36	—	49	28	29	30	29	34	15	02	12	−08	16	17	08	−10	−08	11
Motivation for Vocational Achievement I (7)	25	24	31	32	31	44	—	39	30	41	38	43	16	07	12	−15	18	19	03	−22	−11	18
Motivation for Vocational Achievement II (8)	22	06	18	11	25	18	27	—	10	08	06	15	11	09	13	17	04	09	11	−23	−07	−01
Practical Reasoning																						
Map Reading (9)	35	21	20	13	41	18	26	15	—	62	62	10	09	−18	09	−23	16	10	−09	02	−05	−06
Zip Coding (10)	56	21	20	28	61	33	34	18	41	—	69	29	16	−02	09	−29	24	19	02	−03	−02	07
File Card Sorting (11)[a]	—	—	—	—	—	—	—	—	—	—	—	26	17	−03	03	−27	19	24	02	−14	−12	−04
Enrollee Rating (12)	06	07	28	01	−03	26	26	21	09	06	—	—	10	19	23	13	22	29	16	−18	00	16
Vocational Interest																						
Clerical (13)	14	04	23	23	27	20	48	27	13	24	—	13	—	42	33	−01	52	52	40	05	21	12
Service (14)	15	−03	11	12	09	14	32	13	17	13	—	07	25	—	21	26	46	47	57	04	21	18
Technical (15)	−22	−17	−01	−09	−21	−13	10	13	−10	−22	—	05	11	43	—	30	51	41	37	−04	07	10
Outdoor (16)	−15	−10	−03	−11	−20	−14	−10	15	−15	−20	—	10	−07	19	59	—	12	02	19	−11	01	−10
Science (17)	04	−13	15	20	04	02	19	32	01	−00	—	14	37	39	44	28	—	56	43	−01	11	13
Business (18)	−04	−02	21	05	−09	03	29	07	00	−05	—	15	54	28	45	18	41	—	47	−17	01	26
Aesthetic (19)	−06	−08	03	−03	−07	−02	12	09	−01	−12	—	−07	20	47	57	27	38	41	—	−07	07	19
Vocational Aspirations (20)	05	07	−17	−01	−07	−10	−17	−07	−05	13	—	−01	−14	−08	−18	−08	−15	−22	−22	—	36	02
Vocational Plans (21)	−02	−06	04	−06	04	−05	−03	−02	−03	07	—	00	11	01	−14	−17	11	17	−10	29	—	−05
VA. VP Difference (22)	15	05	−06	04	05	19	05	11	08	−00	—	−10	07	02	07	−03	18	−01	13	−22	−10	—

Note —Loadings for males appear above the diagonal and for females, below the diagonal. Decimal points have been omitted.

[a] Eliminated from analysis for female sample

TABLE 7

ROTATED FACTOR LOADINGS MALE AND FEMALE SAMPLES

Measure	Factors								
	Males				Females				
	I Intellectual Skills	II Positive Attitudes	III General Vocational Interest	IV Negative Self-Image	I Intellectual Skills	II Positive Attitudes	III General Vocational Interest	IV Social Conformity	V High Level Occupational Interest and Motivation
Job Knowledge	.62	.39	−.04	−.15	.76	−.06	−.05	.17	.00
Attitude toward Authority	.29	.52	−.05	.30	.06	.14	−.08	.78	−.11
Self-Esteem	−.05	.70	−.04	−.28	.15	.72	.01	.02	.17
Deferred Gratification	.29	.28	.03	.04	.24	−.10	.04	.75	.14
Job Seeking	.79	.26	.12	−.08	.81	.01	−.10	.12	.10
Job Holding	.31	.70	.05	.04	.29	.44	−.11	.48	.05
Motivation for Vocational Achievement I	.39	.47	.10	−.13	.33	.44	.21	.39	.30
Motivation for Vocational Achievement II	.04	.13	.07	.19	.23	.26	.09	.01	.04
Practical Reasoning									
Map Reading	.82	.02	−.03	.12	.68	.20	.08	.01	−.07
Zip Code	.80	.11	.12	−.10	.79	.10	−.07	.17	.09
File Card	.83	.07	.09	−.04					
Enrollee Rating	.15	.67	.24	.12	−.06	.73	.00	.05	−.03
Vocational Interest									
Clerical	.19	−.02	.70	−.16	.19	.21	.21	.15	.69
Service	−.19	.05	.73	.01	.27	.05	.71	.07	.13
Technology	.01	.17	.60	.33	−.23	.02	.83	−.09	−.03
Outdoor	−.39	.02	.23	.73	−.26	.06	.59	−.08	−.34
Science	.24	.06	.76	.08	−.01	.01	.50	.02	.40
Business	.19	.10	.78	−.11	−.14	.23	.51	.03	.60
Aesthetic	−.09	−.02	.75	.08	−.02	−.09	.76	−.03	.14
Vocational Aspirations	−.03	−.13	−.04	−.07	.03	−.21	−.18	.10	.00
Vocational Plans	−.03	.07	.17	.05	−.01	−.07	−.19	−.15	.65
VA:VP Difference	−.19	.29	.39	−.53	.04	−.27	−.05	.13	.05
Counselor Ratings	.28	.15	.04	.06	.23	.17	.12	.15	−.06
Work Supervisor Ratings	.40	.14	.02	.14	.11	.27	−.07	.00	−.11

Note.—Loadings of .25 or greater which are underlined are used for purposes of factor interpretation.

conventional measures. Prospects for success appear promising when an attempt is made to overcome possible sources of bias in customary test content, format, and administration through use of materials or techniques that bear greater relevance to the academic capabilities and cultural experience of the disadvantaged. The utility of the present measures is suggested by their item characteristics, levels of reliability, concurrent validity with rating criteria, the logic in the underlying dimensions of the battery, and in a degree of concurrence between present research findings and prior results with disadvantaged adolescent groups.

A major and logical finding is the differential role of intellectual and attitudinal factors for the male and female trainee groups. Among females, only an attitudinal dimension is weighted on a rating proficiency criterion. For males, it is the intellectual components that tend to be more heavily weighted in judgments of performance. This is generally consonant with the body of research results concerning adjustment to the educational setting that indicate a major influence of social conformity or "compliance" in the evaluation of female performance (Anastasi & Foley, 1949, ch. 19). It should be noted, however, that there were several specific attitude scales, involving attitudes toward employment and toward the work-training program, that showed equivalent potential value for use with either male or female trainees.

Vocational status or prestige appears to have a pervading influence on occupational perceptions and expectations of adolescents from a poverty-level background—especially for males. This was found not only in judgments of the desirability of particular occupations and their close approximation to status rankings of the general population, but also in the patterns of acceptance of work-related tasks and a status effect that shapes vocational desires in relation to expectations.

A number of the measures used in the battery and the techniques used in their design can be considered to have potential value. But, the results in no way confer a "formal status" on such measures that would justify their application as guidance, selection, or placement instruments. More extensive research investigations with larger samples, in order to provide suitable normative data and development of more varied and longer range criteria, are necessary to arrive at professionally defensible tools for studying disadvantaged groups. The ultimate goal should be one of substituting a standard set of appraisal devices with reasonably understood characteristics for the continuing spate of "ad lib" or informal scales and questionnaires usually employed. Inconsistencies in research results with disadvantaged groups seem inevitable if the findings may be more attributable to measurement inadequacies than to behavioral variations of the population being evaluated.

Admittedly, measures tailored to a disadvantaged subculture do not avoid the paradox inherent in comparing Negro-white, advantaged-disadvantaged, high- and low-SES groups, etc., on test scores that may possess widely different properties (and unspecific "biases") for the groups being compared. Ultimately more meaningful comparisons between such subcultural groups and the comparability of tests must hinge upon the development of long range, comparable criteria from group to group. Nevertheless, the results of the present study hold at least some promise for the application of needed assessment devices *within* culturally and economically deprived populations.

REFERENCES

AMOS, W. E. Job adjustment problems of delinquent minority group youth. *Vocational Guidance Quarterly*, 1964, **13**, 87–90.

ANASTASI, A. Some implications of cultural factors for test construction. *Proceedings of the 1949 Invitational Conference on Testing Problems*. Princeton, N.J.: Educational Testing Service. 1950.

ANASTASI, A. Culture fair testing. *Educational Digest*, 1965, **30**, 9–11.

ANASTASI, A., & FOLEY, J. P., JR. *Differential psychology*. (Rev. ed.) New York: MacMillan, 1949.

BEMIS, S. E. Occupational validity of the general aptitude test battery. *Journal of Applied Psychology*, 1968, **52**, 240–244.

BERNBERG, R. E. *Human relations inventory*. Chicago: Psychometric Affiliates, 1954.

CLARK, C. A. The use of separate answer sheets in testing slow-learning pupils. *Journal of Educational Measurement*, 1968, **5**, 61–64.

CLEARY, T. A., & HILTON, T. L. An investigation of item bias. *Educational and Psychological Measurement*, 1968, **28**, 61–75.

DUNCAN, O., HATT, P., & NORTH, C. *Occupations and social status*. New York: Free Press of Glencoe, 1961.

EELS, K., DAVIS, A., HAVIGHURST, R., HERRICK, V., & TYLER, R. *Intelligence and cultural differences*. Chicago: University of Chicago Press, 1951.

EMPEY, L. T. Social class and occupational aspiration: A comparison of absolute and relative measurement. *American Sociological Review*, 1956, **21**, 703–709.

GEIST, H. *The Geist picture interest inventory:*

Male and female (Manual). Beverly Hills, California: Western Psychological Services, 1964.

GORDON, E. W. A question of culture. *American Child*, 1963, **45**, 11–14.

KAISER, H. F. The Varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 1958, **23**, 187–200.

KARP, J. M., & SIGEL, I. Psychoeducational appraisal of disadvantaged children. *Review of Educational Research*, 1965, **35**, 401–412.

KIRKPATRICK, J., EWEN, R., BARRETT, R., & KATZELL, R. Differential selection among applicants from different socioeconomic or ethnic backgrounds. (Final Report) New York: New York University, Department of Psychology, May 1967.

KNAPP, R. R., GRANT, B., & DEMOS, G. D. *California Occupational Preference Survey*. San Diego: Educational and Industrial Testing Service, 1966.

LENNON, R. T. *Testing and the culturally disadvantaged child*. New York: Harcourt, Brace & World, 1964.

LOPEZ, F. M., JR. Current problems in test performance of job applicants: I. *Personnel Psychology*, 1966, **19**, 10–18.

POTTHOFF, R. Statistical aspects of the problem of biases in psychological tests. (Mimeo Series No. 479) Chapel Hill: University of North Carolina, Department of Statistics, May 1966.

ROSENZWEIG, S. *Picture frustration study*. St. Louis, Mo.: Author, 1948.

SCHNEIDER, L., & LYSGAARD, S. The deferred gratification pattern: A preliminary study. *American Sociological Review*, 1953, **18**, 142–149.

SOCIETY FOR THE PSYCHOLOGICAL STUDY OF SOCIAL ISSUES. Guidelines for testing minority group children. *Journal of Social Issues*, 1964, **20**, 127–145.

STEPHENSON, R. Occupational aspirations and plans of 443 ninth graders. *Journal of Educational Research*, 1955, **49**, 27–35.

STRAUS, M. Deferred gratification, social class, and the achievement syndrome. *American Sociological Review*, 1962, **27**, 326–335.

THORNDIKE, E. L., & LORGE, I. *The teacher's word book of 30,000 words*. New York: Columbia University, Teacher's College, 1944.

UNITED STATES DEPARTMENT OF LABOR, MANPOWER ADMINISTRATION. *The neighborhood youth corps program*. Washington, D. C.: United States Government Printing Office, 1965.

WESMAN, A. G. Intelligent testing. *American Psychologist*, 1968, **4**, 267–274.

WISPÉ, L. A psychologist looks at motivational problems in training. *Occupational Outlook Quarterly*, **9**, 1965, 13–15.

(Received June 9, 1969)