VOLUME 9 NUMBER 8 JANUARY 1978

FOCUS ON EXCEPTIONAL CHILDREN

Nondiscriminatory Testing Of Minority and Exceptional Children

Gordon Alley and Carol Foster

"When is Washington's Birthday?" If you are of the majority culture, you probably would answer "February 22." This response would be scored correct according to the Wechsler Adult Intelligence Scale scoring criteria (Wechsler, 1955, p. 34). If, however, you were to respond "April 5," your response would be scored as incorrect, using that scoring criteria. The implication is that you must associate *George* rather than *Booker T.* with *Washington.* This item could be classified as culture-biased (example provided by Williams, 1974, p. 16).

In like manner, a visually handicapped child might be asked, "What should you do if you see a train approaching a broken track?" This handicapped child might have to reorient his or her learning strategy and problem solving skills to parallel those of a sighted group in giving a "correct" response such as, "Wave a handkerchief."

These examples are not a condemnation of the Wechsler measures. Similar examples are found on most measures. The examples are representative of items that have caused at least one group of psychologists to advocate calling a moratorium on administration of conventional psychological tests to minority groups (Position Statement of the Association of Black Psychologists, adopted at a meeting of the Association in Washington, DC, August 1969).

Concurrent to this position statement, the judicial branch of the U.S. Government was hearing the *San Antonio Independent School District* v. *Rodriquez* case (1973). The ruling included an explicit statement that traditional assessment measures are unsatisfactory when used as a side measure to identify exceptional children who represent minority groups.

The Civil Rights Act of 1964, Title VI, require(s) that "there be no discrimination on the basis of race, color, or national origin in the operation of any programs benefiting from Federal financial assistance" (Memorandum for Chief State School Officers and Local School District Superintendents, DHEW, Washington, DC, August 1975, p. 1). Other legislation (Title IX of the Civil Rights Act, 1964 (1972); PL 93-380, which amended Part B of the Education of the Handicapped Act (PL 91-230) (1969); and PL 94-142 (Education for All Handicapped Children Act of 1975) followed.

Gordon Alley, Professor, and Carol Foster, Research Associate, are staff members of the Department of Special Education, University of Kansas.

Title IX of the Civil Rights Act requires that no program shall discriminate against a person because of sex. Overinclusion or underinclusion of children of either sex in any special program category can suggest non-compliance. In addition, using criteria or methods of referral, placement, or treatment which in effect discriminate because of sex also can constitute non-compliance.

Public Law 93-230 includes specific standards relating to testing and assessment. The following two standards are included as part of the requirements for state plans:

1. Failure to adopt and implement procedures to ensure that test materials and other assessment devices used to identify, classify, and place exceptional children are selected and administered in a manner which is nondiscriminatory in its impact on children of any race, color, national origin or sex. Such testing evaluation materials and procedures must be equally appropriate for children of all racial and ethnic groups being considered for placement in special education classes. In that regard procedures and tests must be used which measure and evaluate equally well all significant factors related to the learning process, including but not limited to consideration of sensorimotor, physical, socio-cultural and intellectual development, as well as adaptive behavior. Adaptive behavior is the effectiveness or degree with which the individual meets the standards of personal independence and social responsibility expected of her or his age and cultural group. Accordingly, where present testing and evaluation materials and procedures have an adverse impact on members of a particular race, national origin, or sex, additional or substitute materials and procedures which do not have such an adverse impact must be employed before placing such children in a special education program.
2. Failure to assess individually each student's needs and assign her or him to a program designed to meet those individually identified needs (p.3).

The most explicit standards to assure nondiscriminatory testing of exceptional children are included in the regulations provided for compliance of PL 94-142. Both the evaluation procedures and placement assessment are covered in these regulations:

121a532 Evaluation procedures.
State and local educational agencies shall insure, at a minimum, that:
(a) Tests and other evaluation materials:
(1) Are provided and administered in the child's native language or other mode of communication, unless it is clearly not feasible to do so;
(2) Have been validated for the specific purpose for which they are used; and
(3) Are administered by trained personnel in conformance with the instructions provided by their producer;
(b) Tests and other evaluation materials include those tailored to assess specific areas of educational need and not merely those which are designed to provide a single general intelligence quotient;
(c) Tests are selected and administered so as best to insure that when a test is administered to a child with impaired sensory, manual, or speaking skills, the test results accurately reflect the child's aptitude or achievement level or whatever other factors the test purports to measure, rather than reflecting the child's impaired sensory, manual, or speaking skills (except where those skills are the factors which the test purports to measure);
(d) No single procedure is used as the sole criterion for determining an appropriate program for a child, and
(e) The evaluation is made by a multidisciplinary team or group of persons, including at least one teacher or other specialist with knowledge in the area of suspected disability.
(f) The child is assessed in all areas related to the suspected disability, including where appropriate, health, vision, hearing, social and emotional status, general intelligence, academic performance, communicative status and motor abilities (Federal Register, August 23, 1977, pp. 42496-42497).

CURRENT APPROACHES TO NONDISCRIMINATORY TESTING

To evaluate the various alternatives which have been suggested to be nondiscriminatory, one initially must present specific criteria for such an evaluation. First, we offer the following definition of a nondicriminatory measure to be used in our evaluative task:

FOCUS ON EXCEPTIONAL CHILDREN is published monthly except June, July, and August as a service to teachers, special educators, curriculum specialists, administrators, and those concerned with the special education of exceptional children. This journal is abstracted and indexed in *Exceptional Child Education Abstracts,* and is also available in microfilm from Xerox University Microfilms, Ann Arbor, Michigan. Subscription rates, $10.00 per year. Printed in the United States of America. Second class postage is paid at Denver, Colorado.

Executive and Editorial Office
6635 East Villanova Place
Denver, Colorado 80222
Telephone (303) 757-2579

A nondiscriminatory measure is one which results in similar performance distributions across cultural groups. These cultural groups may differ with respect to any or all of the following:

1. language/dialect
2. value system
3. information
4. learning strategies.

Regardless of the purpose of a test *or its validity for that purpose*, a test should result in distributions that are statistically equivalent across the groups tested in order for it to be considered nondiscriminatory for those groups. If different groups' performances result in different distributions, the test discriminates among groups.

For some variables, such as height and weight, the measures that have been used have been demonstrated to have such high validity, reliability, and little error in measurement that differences occurring among populations are considered *real* rather than flaws in the measurement device.

For other variables, especially psychological ones such as intelligence and achievement, the measures have had much lower validity, reliability, and fairly significant measurement error, resulting in less confidence. When minority groups produce performance distributions that vary from the majority group distribution, a strong tendency is to assume that the differences are in precision of the measurement and that these differences are not *real* between the groups. Because of the many problems in measuring intelligence and achievement, this seems to be a most parsimonious explanation.

If we accept that intelligence and achievement tests are discriminatory across various cultural groups, we must then ask, "How can we measure intelligence and achievement in all of those groups so that we are not discriminatory?"

Using the criteria described above, one can evaluate testing and assessment procedures that have been popular in their use as nondiscriminatory methods. Four popular procedures have been advocated; namely:

1. Translating traditional tests from the majority language directly to the minority language;
2. Norming traditional tests on specific groups of minorities;
3. Using a minority examiner to test minority children; and
4. Identifying majority group competencies required for minority group children to survive in the majority culture, then evaluating the minority child's achievement of these competencies, and, finally, teaching the child the unattained competencies.

All of these procedures appear satisfactory when given cursory attention. A critical evaluation of these procedures however, yields information that may discourage their use in approaching nondiscriminatory testing.

First, consider procedure 1, literally translating the Stanford-Binet, Raven, Peabody Picture Vocabulary Test, etc. into the native language of a minority group, say, Spanish-speaking children. This procedure appears to be a simple and efficient method of equating the language difference of majority group children with minority group children. In fact, PL 94-142 specifically states that this procedure meets one of the requirements of nondiscriminatory testing. But language differences are *not* equated by this procedure when one considers the complex language idioms, colloquialisms, words with multiple meanings, and words of similar but not identical meanings that characterize all languages.

Garcia (1976) provided an excellent example of problems associated with literal translation. On one test, a question in English contains the words, "hot dog." This term translated literally into Spanish means "a female dog in heat." The structure of the translated statement was changed to the extent that the children did not know how to respond to such a nonsensical statement. In addition to the above problem, the translation procedure does not equate for the differing cultural information, learning strategies, and value systems when the test items are not changed to reflect these factors as they occur among the different groups.

Procedure 2 suggested to fulfill nondiscriminatory testing evaluation criteria appears more satisfactory. When normative data are available on children of a minority group, this permits better comparisons of one minority child's performance to the performance of other children of the same minority group. Several alternatives of this procedure have been suggested for use (Mercer, 1973; Thorndike, 1974; and Williams, 1974).

Thorndike (1974) stated that some test users consider a measure to be nondiscriminatory if there is no difference in mean scores and/or variability of two culturally different groups on a test, and/or if the regression equation developed on one cultural group neither overpredicts nor underpredicts another cultural group's per-

formance on the test. He argues that difference in means and/or variance per se does not constitute discriminatory testing; rather, "one must examine the correlates of those differences" (p.37). On one hand, if there is *no relationship* between a *criterion* variable (e.g., highest grade attended in school) and *predictor* variable, (e.g., score on a test), one has no basis for judging the nondiscriminatory qualities of the measure. Conversely, if the relationship of the two variables is established, the test score is obviously discriminatory for the measure defined—i.e., low scores on test are related to highest grade attended.

Thorndike provides an excellent example of a discriminatory intelligence test. A situation exists in which only students obtaining IQ scores of 100 or above were selected for clerical positions; the test is obviously unfair to minority groups.

> . . .if it were found that, given certain conditions of adaptive and remedial instruction, a group of culturally deprived youngsters with a mean IQ of 85 (predictor) could be brought to the same level of proficiency in a clerical position (criterion) that was displayed by an unselected sample with an IQ of 100. . . (p.40).

Thus, the group IQ score discriminates fairly between the two groups of youngsters, given no job training is provided. If job training is provided to the selected applicant(s), the group IQ score discriminates unfairly between the two groups of applicants.

Thorndike provides two definitions of nondiscriminatory use of tests. The first is appropriate to this discussion. It provides that both the majority group and the minority group have the same opportunity for selection to special services (criterion) as would be represented by the proportion of the group falling below critical score on the test of academic performance in the regular classroom (predictor). This definition can be implemented by setting a critical score on the achievement test (predictor) for the majority group, based on the percentage found to benefit from special services in the past (criterion). *Independently*, one must set a critical score on the achievement test for the minority group, based on the percentage found to benefit from special services in the past (criterion).

This procedure requires that one have specific information related to performances of the minority group on both the predictor variable (achievement test) *and* the criterion variable (beneficial effects of special services). This procedure, however, does not consider intra-minority group differences. Garcia's (1976) statement is particularly relevant here:

> Make no assumptions about the bilingualism of Mexican-American or other linguistically-different students. *Some* may be fully literate in two languages. *Others* may speak only English (p.2).[1]

Mercer (1973) has attempted to meet this criticism. She suggested the use of pluralistic norms for interpreting the meaning of a test score (e.g., I.Q. to predict mental retardation status). She stated that it is not possible to consider blacks or Mexican-Americans as homogeneous social categories or to ". . .hold sociocultural factors strictly constant by controlling only for ethnic group" (p.248).

Pluralistic norms ". . .evaluate the. . .(performance) of a person only in relation to others from similar sociocultural backgrounds" (p.248).[2] She provided an excellent example of use of the pluralistic approach:

> . . .if he is a Mexican-American child and manages to achieve 75 on an intelligence test when he comes from an overcrowded, Spanish-speaking home in which the father has less than an eighth grade education and was reared in a rural area, and his mother does not expect him to finish high school, he would be diagnosed as having normal ability (p.249).

With reference to the pluralistic approach, the reason this child is classified as normal and not to be considered for special education services is that he scored within one standard deviation of his sociocultural modality group. His low score on the intelligence test reflects his lack of opportunities rather than a general learning deficit. Mercer (1973) suggested that the pluralistic approach may be used with other minority groups. The only requirement is that the predictor test score is interpreted within the framework of each minority modal grouping. One also might note that ". . .nothing happens to Anglo rates when pluralistic norms are applied" (p.254).

This second procedure of obtaining pluralist normative data on children of minority groups, however, contains several severe limitations. Thorndike (1974) lists three limitations:

— relevancy
— reliability
— bias.

Relevancy according to Thorndike in describing criterion variability which has been unaccounted for by the test, is only partially attained on tests of either prediction or criterion. For example, achievement test measures are not constructed to measure *all* aspects of school per-

[1]Italics added for emphasis.
[2]Information in parenthesis added by authors for contextual clarity.

formance. Therefore, one can only guess whether the unmeasured aspects of school performance will enhance or handicap a minority child. In addition, no school achievement test accounts for 100 percent of the variance of actual school achievement. Even if the test accounts for 81 percent of the variance (r = .90), which is generally higher than typically occurs in practice, one cannot account for 20 percent of the variance. Thorndike aptly states, "It becomes impossible to be sure what adjustment in critical score, if any, is appropriate for minority group members" (p.45) to yield a true measure of achievement. This same argument holds for majority group members.

If the measure is unreliable, one must make an estimate of the true criterion performance. Thorndike provides a rationale and statistical procedure to obtain an estimation of the true criterion difference based on means, standard deviations, and reliability coefficient of both the majority and minority groups to provide statistical fairness. One then must decide if this statistical fairness provides a socially fair test. Resolution has not been made to the satisfaction of the present authors.

Finally, "if the criterion measure is itself biased in an unknown direction or degree, no rational procedure can be set up for the 'fair' use of the test" (p.44). That is, the criterion performance must mean the same thing to both majority and minority children. School success is a good example of criterion performances. To some children, in both the majority and minority groups, school success is not measured by academic performances (e.g., grade in English composition), but it may have meaning in vocational education performance (e.g. grade in automotive repair). For these children, a criterion test of only academic performance is biased and discriminatory to their value systems.

Williams (1974) provided the reader with the lead question of this article. He considered the response "George Washington" as a biased estimate of the black child's knowledge. Williams constructed the Black Intelligence Test Counterbalanced for Honkies (BITCH) (Williams, 1974). All items were obtained from black culture, and the test "...is biased in favor of black people" (p.16). He developed this measure to emphasize the need to revalidate conventional tests to responses of black persons on "white-oriented tests" (p.17).

Alley (1976) provided a fourth limitation of the procedure of obtaining normative data on children of minority groups on conventional measures, stating that the procedure solidifies the status quo of minority children. Using Mercer's (1973) example of the Mexican-American child scoring an IQ of 75, Mercer interpreted this performance to be normal functioning for the sociocultural modality group of which he was a member; i.e., children from that particular sociocultural modality group or any other such group would *"always perform poorly"* and that no attempt would be made to find the child's areas of strength or to search for defects in the test.

On the basis of the four criteria of equality, this second procedure tangentially considered the information of the minority group child from the majority group's standards (Mercer, 1973). It does not consider the child's value system as it relates to the criterion performance (Thorndike, 1974). This procedure also does not consider either the language structure of the items or the learning strategies of the minority children (Garcia, 1976).

Williams (1974) contends that ". . .there is a white psyche and a black psyche" (p.17) and that "white psychiatrists, white psychologists, white social workers, and other white mental health workers cannot successfully treat the black psyche" (p.18). The reason that white mental health persons cannot treat most black persons is that they do not understand what Williams calls "niggerosis"; i.e., being black, being called "nigger" and being told you are unintelligent when you are intelligent. His solution is to train more black professionals to treat the black psyche. This contention and solution also are advocated by Garcia (1974) for bilingual children in his suggesting this third procedure to obtain nondiscriminatory test results. He recommends:

> Be skeptical about utilization of standard diagnostic instruments when used to identify the learning behaviors and capabilities of bilinguals. Instead, utilize bilingual clinicians to assist in the identification process (p.3).

Most persons would agree that many white teachers and other white professionals associated with minority group children neither understand nor communicate well with these students or clients.

The solution of providing a minority group examiner to administer a test to a minority child, however, is simplistic. Three conditions have been overlooked by persons advocating this procedure. First, the attitudes of one person toward another may reflect social class differences to a greater extent than racial or ethnic differences. In Wagner (1972), Clark provides one reason not to use some minority group examiners to test minority group children:

> Many of today's scholars and teachers came from (culturally deprived)[3] backgrounds. Many of these same individuals, however, when confronted with students whose present social and economic predicament is not unlike their own past tend to react negatively to them, possibly to escape the painful memory of their own prior lower status (p.131).

Wagner (1972) provides a second reason, characterizing some persons' attitudes as, in essence:

> "I came from a neighborhood like this, and I pulled myself up without all the help which is being provided you; you can pull yourself up too," and (then they) drive away to their suburban homes (p.440).

The present writers believe that the key word of persons advocating this third procedure is *empathy*.

Third, is empathy enough? The authors would suggest that choice of examiner, whether the examiner represents any one of several sociocultural modality groups, racial or ethnic minorities, is not enough to assure nondiscriminatory testing. Excellent as this procedure may appear to be, the examiner must be provided with more than the conventional, culturally-biased tests. Even in the hands of the most competent and empathic examiner, regardless of his or her group membership, the minority child cannot display competence on these conventional tests. Thus, the third procedure will obtain nondiscriminatory results only if the selected minority examiner is provided with alternative measures that more appropriately evaluate the child's competence.

This third procedure *assists* in obtaining nondiscriminating testing if the examiner possesses similar language, value system, cultural information, and learning strategies as those of the child. If the examiner is to administer the *conventional* discriminatory tests, however, administration of the test will be frustrating to both examiner and child.

The fourth and final procedure to obtain nondiscriminatory testing is to identify and teach competencies required to survive in the majority culture. This procedure is analogous to teaching "cram courses" to servicemen who were selected for duty in countries with a language and socioculture value systems different from their own. The courses did teach the servicemen minimal language competence, but generally did not consider the differences in value systems, cultural information, and learning strategies of the people. The result was that these serviceman could ask some questions and minimally converse, but they required more experiences to become integrated into the sociocultural milieu of the country.

The experience of one of the present writers in Japan and Korea permitted him to witness the integration of some servicemen into these countries' sociocultural systems. Many of the servicemen who become integrated into the societies were neither completely knowledgeable of the country nor proficient in the country's language, but they shared common experiences with the people and did not permit cultural differences to alienate them from the people. This personal experience has been corroborated by reports from Peace Corps volunteers who reported sustaining friendships with persons, irrespective of language or sociocultural differences, when they shared experiences common to each other.

The point inherent in the above two reports is that the fourth procedure arbitrarily and explicitly places a higher value on the majority group's language and cultural information. Implicit to the procedure, but explicit to the child, is that the language, value system, cultural information, and learning strategies of the minority group are inferior. Stacker (1967) suggests that the result of this procedure on the Mexican-American student is ". . .that his culture is no good and therefore assumes that *he* is no good and grows up with this attitude, completely denying his own culture and value system" (p.439). Coleman (in Wagner, 1966) reported that his committee found that parents stopped reinforcing the child's positive attitudes toward school as instruction content became irrelevant to the parents. Williams (1974) stated that the black child has had to leave both verbal and cultural skills outside the classroom because these skills are not rewarded in the middle-class classroom.

We have questioned the conventional use of nondiscriminatory or "fair" testing. Thorndike has stated that this conventional usage ". . .is based on predicted criterion performance (*survival in the majority sociocultural system is set by one who is almost always the test constructor*)[4] at some level or standard of predicted criterion performance as the requirement for acceptance, and would apply this *majority group* standard to both majority and minority groups" (p.43).

All four procedures to obtain nondiscriminatory testing have serious flaws inherent in their rationale and/or use. Two options remain: First, one could follow the lead of Williams and the Association of Black Psychologists in calling for a moratorium on conventional psychological tests until truly nondiscriminatory tests are available for use; or, second, one might seek innovations that may prove more productive in the search for nondiscriminatory tests and testing.

[3]Information in parenthesis is added by author for contextual clarity.

[4]Parenthesis are the authors.

Messick and Anderson (1974) provide the consequences of choosing the first option. They suggest that examiners then would turn to subjective appraisals in which the likelihood of discrimination of minority groups inevitably would increase. A second consequence would be that data generally would be gathered unsystematically. A third consequence would be increased parochialism without benefit of regional or national norms. The three consequences are not of equal importance, in these authors' opinion, but the sum and substance of the results would suggest that one choose the second option.

ADVANCES IN NONDISCRIMINATORY TESTING

Recently, several innovative assessment procedures have yielded optimistic results that may have provided direction toward the development of nondiscriminatory tests and testing practices. These assessment procedures have sought to focus on the competencies of exceptional children rather than on their deficits, and appear to result in more equal performances of sociocultural subgroups of the majority and minority groups.

The first procedure is *clinical teaching* (Lerner, 1976). Clinical teaching does not tap what a child has learned, but rather provides experiences for children to actively involve themselves in problem solving. Budoff and Hutton (1972) provide the learning potential procedure to probe for competencies among minority group members who are considered to be exceptional. Budoff uses a testing measure that is less culturally-biased and language-oriented than traditional measures of problem solving—Raven's Progressive Matrices. He has found that some children viewed as poor problem solvers apparently have not profited from majority group cultural information. Using Raven's Progressive Matrices, 50 percent of those children scored at or above the average range when the task was changed to a less majority-biased learning format.

Even more dramatic, Budoff and Hutton found that if they provided only one hour of structured experiences in problem solving to children who initially scored low on the Raven's, 50 percent of these low performing children scored at the 50th percentile or above on the posttest given after this short training. The latter group, "gainers," overrepresented minority groups. Platt (1976) and Swanson (1976) found a similar result with learning disabled children. This is a major assessment breakthrough in the identification of exceptional children, particularly of exceptional children representing minority groups.

Working independently of Budoff, Flavell (1975) has been studying the learning strategies of preschool and primary grade children. He found that nonhandicapped children who could state strategies to retrieve information performed better at memory tasks than those who could verbalize no strategy to retrieve this information. Anderson (Anderson & Alley, 1977), using a similar approach with a problem solving discrimination task and using matched age mentally retarded and normally functioning children, found that the label of MR or nonhandicapped was not as important to success on this task as was the ability to verbalize a strategy to solve the problem. These findings suggest a second approach to assessing the competencies of exceptional children.

A third procedure that has been studied for a long time but has been given little practical attention by school psychologists, counselors, and teachers is the importance of the child's adaptive behavior. Adaptive behavior refers to the extent to which a child meets the cultural and societal demands in her or his environment. (Mercer, 1973). Such activities as self-help skills, language, personal and social relationships, and vocational competencies are measured with academic competence. Using this broader base of information, one can judge the exceptional child's competencies in total living skills rather than only a narrow cognitive area (e.g., knowledge of fractions).

Nondiscriminatory Testing with the Severely/Profoundly Handicapped

If we examine the issue of discriminatory testing with regard to the severely and profoundly handicapped, we must ask if the individuals, when grouped according to cultural backgrounds, produce different distributions. To the authors' knowledge, this question has not been answered. On intelligence tests for the general population, few, if any, included severely handicapped individuals in their population for developing norms, and they measure grossly at the low end of the scale. Nevertheless, a high correlation ($r = .88$) exists between individuals being classified at various levels of intelligence tests and a clinical psychologists's classification of

adaptive behavior (Leland, Nihira, Foster, Shellhaas, & Kagin, 1966). Therefore, although these tests may not make fine discriminations (at least if one accepts a clinical classification as a useful referent for validity), they do seem to be related to general functioning.

The question of whether they make these distinctions equally well independent of cultural background is still unclear. Mercer (1973) found that the tests had a heavy cultural bias when dealing with mildly retarded individuals. This discussion concludes with the authors' observation that this is an empirical question which has not been adequately approached, but that limited evidence seems to indicate that it is not seen as much of a problem.

Sailor and Horner (1976) examined the same issue with regard to tests designed specifically for the severely and profoundly handicapped. Their focus is on global educational assessment devices which can be used to assess individuals and describe their current skills and skill deficits. Older tests had the goal of determining how "retarded" an individual was but, with the change in emphasis to educational programs, the newer assessment tools were used more to prescribe the individual's educational prescriptive program.

The scales reviewed by Sailor and Horner include the Vineland Social Maturity Scale (Doll, 1947), the Cain-Levine Social Competency Scale (Cain, Levine, & Elzey, 1963), the AAMD Adaptive Behavior Scale (Nihira, Foster, Shellhaas, & Leland, 1974), the Balthazar Scales of Adaptive Behavior (Balthazar, 1971a, b, c, d; 1973), the Camelot Behavioral Checklist (Foster, 1974), the Portage Project Checklist (Shearer, Billingsley, Frohman, Hilliard, Johnson & Shearer, 1970), the Pennsylvania Training Model: Individual Assessment Guide (Somerton & Turner, 1975), the APT (Pennhurst Assessment/Program Tool, 1976), and the TARC System (Sailor & Mix, 1975). Many other assessment devices are aimed at this population, but these provide a good overview and will be used as the basis for this discussion.

One way in which most of these assessment devices differ from standardized intelligence tests is that they either consist of checklists or involve direct observation of individuals performing specific behaviors. The interest of the evaluator is in that behavior per se, and not in that behavior's theoretical relationship to some global characteristic for which it is assumed to be a measure, such as intelligence. Here, the emphasis is on identifying whether or not an individual has a certain skill. If she or he lacks that skill, it needs to be taught. If she or he possesses that skill, no further inference is necessary about intelligence or functioning level—the conclusion is merely that this skill does not have to be included in an educational program for this individual.

Assessment of procedures with severely handicapped individuals appears at first glance to be one area in which tests and evaluation instruments are relatively nondiscriminatory; i.e., one has little reason to assume that the effects of one cultural background versus another would produce deviations in the results of the tests large enough to be even noticeable, when compared to the differences that the tests measure which are consistent across almost all cultures. For instance, although cultures may differ widely on the techniques an individual uses to dress or feed her or himself, they all are in fairly close agreement that one should be able to dress or feed oneself.

Because most of the assessment instruments are geared more toward the grosser assessment of skills, variations in skills related to a specific cultural background are of little importance. Individuals across a wide variety of cultural backgrounds would be scored fairly similarly on assessment devices for the severely handicapped, and individuals who scored significantly lower on the devices generally would be considered to have major impairments, regardless of cultural backgrounds. For instance, if an individual has no expressive language and little receptive language, she or he would be considered impaired for adaptation to any culture. Also, little error would be found in diagnosing a deficit at this level, as opposed to assessing someone who has a fair amount of language skills embedded in a culture different from the one of the test or of the test giver.

One characteristic common to all these devices is that they assess over a large number of skills in many different domains. They almost all look at the following areas, at least: self-care, cognitive or academic, language, motor skills, and vocational. Because they examine an individual across such a broad base, and it is generally assumed that skill deficits will occur across all of them, if individuals are low on only one of these areas, they would be subjected to much more extensive assessment before a determination of diagnosis or program is made. This procedure is one of the safeguards against improper diagnosis and misuse of assessment instruments for the severely handicapped.

Another safeguard is that the responsible professionals typically do not depend upon one sample of the individual's performance which can be adversely affected by a large number of variables such as illness, medication,

attitude, etc. Most of the devices (e.g., the Adaptive Behavior Scale, the Vineland, and the Camelot Behavioral Checklist) are scored by someone who knows the individual quite well, and who rates the individual's typical performance. Since these checklists and scales almost always refer to observable behaviors, scoring of them can be done fairly quickly and reliably (e.g., inter-observer reliability for the Adaptive Behavior Scale, pp.71 to 92, for Part I which measures adaptive skills, and for the Camelot Behavioral Checklist, p.93).

Even this assessment is seen as preliminary for checklists which are prescriptive instruments (e.g., the Portage Project Checklist, the Pennsylvania Training Model, and the Camelot Behavioral Checklist). After the checklist is completed, instructions are provided for determining what specific skill training programs the individual should have. The first step of the training is a pre-assessment which involves direct observation of the individual attempting to perform the task in question. If the individual demonstrates the skill, the next skill in the sequence is tested. This procedure continues until the pretest indicates that the individual cannot perform the task. A few checklists (the Balthazar Scales of Adaptive Behavior; the Pennhurst Assessment/Program Tool) utilize direct observation of the individual for completing the entire assessment. This can be done if one is interested in investing a great deal of time in assessment or if one assesses small segments of performance in detail.

One other emphasis of this area of assessment which frees it from being discriminatory is the degree to which most of the instruments avoid labels and static diagnoses. Instead, they emphasize assessment of specific skill deficits, which leads to a specific educational program to promote that skill. Here, the reason for assessment is strictly to determine what skills an individual requires to adapt to the general demands of society, and then proceeds directly to teaching that skill.

In general, then, one can conclude that assessment procedures of severely handicapped individuals are not subject to discrimination on the basis of cultural biases because:

1. They assess behaviors which are generally similar across cultures;
2. They assess the individual in ways that are not influenced by specific temporary states of the individual;
3. They assess a wide range of skills, and individuals who have deficits in only one area would be treated differently from those who lack skills in all areas;
4. They emphasize prescriptive educational programs, sometimes to the exclusion of labels or diagnoses;
5. They depend on direct observation of the individual performing or failing to perform certain skills either during the assessment itself, or as a backup to initial assessment.

These factors combine to produce assessment instruments which are valid for the purpose of discriminating individuals who are severely handicapped from those who are not. Although some problem is presented in discriminating borderline cases, of course, this provides no basis for concluding that the devices are discriminatory along any cultural or ethnic basis.

One problem of these assessment instruments, however, relates to their relative inadequacy to determine specific handicaps which may affect habilitation programs. According to the Bureau of Education for the Handicapped, the term *severely handicapped* refers to:

> From Title 45 Public Welfare
> Chapter 1-Office of Education, Dept. of HEW
> Programs for the Education of the Handicapped
> Part 121-Definitions, General Provisions
>
> "Severely handicapped children" are those who because of the intensity of their physical, mental, or emotional problems, or a combination of such problems, need educational, social, psychological, and medical services beyond those which are traditionally offered by regular or special educational programs, in order to maximize their full potential for useful and meaningful participation in society and for self-fulfillment.
>
> (a) The term includes those children who are classified as seriously emotionally disturbed (including children who are schizophrenic or autistic), profoundly and severely mentally retarded, and those with two or more serious handicapping conditions, such as the mentally-retarded blind and the cerebral palsied deaf.
>
> "Severely handicapped children" (1) may possess severe language and/or perceptual-cognitive deprivations, and evidence normal behaviors such as (i) Failure to respond to pronounced social stimuli, (ii) Self mutilation, (iii) Self-stimulation, (iv) Manifestation of intense and prolonged temper tantrums, and (v) The absence of rudimentary forms of verbal control, and (2) many also have extremely fragile psychological conditions.

Because this definition is so broad, it encompasses many specific disabilities. Individuals applicable to this definition are all similar in that they have extensive skill deficits in many different areas, but an individual who is blind-deaf, is of normal intelligence, and has had a relatively poor education program may appear the same on the assessment instrument as an individual who has

severe retardation and an orthopedic impairment. These two individuals may have identical assessment profiles, and may have the same skills targeted as their next educational goals. Nevertheless, they must be treated quite differently beyond this point. If they are both to be taught skills of self-feeding, one child may need equipment adapted for orthopedic impairments, and might be taught through modeling and verbal instructions; the other may be taught using regular eating utensils, with physical guidance being faded out. For these two individuals, similar results on prescriptive assessment devices should result in totally different programs.

The above example may seem purely academic since few professionals would train an individual who was obviously visually impaired and deaf the same way they would train a severely retarded orthopedically handicapped individual. Handicapping conditions, however, are not always obvious. Procedures for determining intelligence, blindness, deafness, range of motion, language skills, and so on among the severely impaired are not well validated. Procedures for audiometric evaluation of the severely retarded though operant techniques are fairly well standardized (Fulton, 1971), but they are used in few instances; and the technology for assessing other physical attributes (e.g., visual acuity) is still under development (Spellman & DeBriere, 1976).

The impact is that, although assessment devices for the severely impaired do not appear to be discriminatory on cultural or ethnic bases in describing the individual's current ability to adapt to society, they may result in all severely handicapped individuals' being treated in the same way. This may result in inappropriate training strategies or goals to be used if information on the associated handicaps is unavailable. Although some of this technology exists or is under development (Fulton, 1971; Spellman & DeBriere, 1976; Foster & Barnes, 1977), it does not appear to have had the same impact as that of assessment instruments used for prescriptive programming.

CONCLUSIONS

Tests are built to discriminate among individuals — If everyone scored exactly the same, the test would be useless for any purpose. Tests, however, should discriminate *only* along the variables they were designed to measure (e.g., intelligence, achievement), and not along other measures such as cultural background, sex, or value systems.

Although the use of nondiscriminatory test measures is characterized as being problem-wrought for the teacher/psychologist/administrator and the child being evaluated, some reasonable guidelines are suggested for application, along with a list of research questions which must be addressed so the problems do not remain the same or worsen in the future.

The authors make the following recommendations for individuals responsible for administration, interpretation, or decision making on the basis of test results.

Recommendations Related Specifically to PL 94-142 and Nondiscriminatory Testing

1. If one has reason to believe that a test may be discriminatory with a given child, use a large test battery and apply the best results in making any decision based on this testing. Do not make any decisions on the basis of one test alone. *This recommendation relates specifically to Section 121a 532(d).*
2. Validate the results of any evaluations by observations of behavior in natural settings. That is, if a child appears to be below norms on measures of intellectual functioning, does this appear to be substantiated in all other situations? Does the child do better in some classes than could be expected from the test results? Are his or her social skills quite high? Is the child fairly competent in language skills, even though the language is dependent upon cultural background? If routine observation results in one's questioning results of the assessment, assessment results should be examined more closely before making decisions based on them. Also, certain alternative procedures, such as testing by a selected person representative of that minority culture, may prove useful. *This recommendation relates specifically to Section 121a 532(a).*

3. The IQ provides a general estimation of a child's performance. As such, this measure of global functioning has minimal relevance to instructional planning. Measures of specific educational domains are most relevant to educational planning and should be included in the assessment process; educational domains include academic, performance, vision, hearing, social and emotional status, and communication skills. Within each domain are major sub-

domains. The Federal Register (November 29, 1976) suggests major subdomains to be assessed when identifying children with learning disabilities; namely: —Verbal (oral) expression;
—Listening comprehension;
—Written expression;
—Basic reading skills;
—Reading comprehension;
—Mathematics calculation;
—Mathematics reasoning; and
—Spelling.

This recommendation is specifically related to Section 121a532 (b) and (f).

4. In evaluating a child with a suspected disability, PL 94-142 (Section 121a532(e)) *requires* the use of a multidisciplinary team or group, which must *include* either a teacher or *specialist* with knowledge in the area of the suspected disability. The authors recommend this group process with several reservations. A group is superior to an individual when accomplishing intellectual tasks (Davis, 1969). This superiority is based upon the group providing a higher frequency of judgments, more opportunities to correct errors, and a greater chance for one of the group members to possess the skills to solve a complicated task. PL 94-142 has contributed to a high probability of group success by stating that the composition must include a specialist (Middlebrook, 1974). But several reservations are in order:
   a. The group may subject members to intense social pressure to conform and subvert individual efforts—i.e., "group-think" (Janis, 1972), "risky shift phenomenon" (Tubbs & Moss, 1974), and group cohesiveness.
   b. The task may require a sophisticated and high degree of coordination and organization. If so, an individual may be more effective than a group.
   c. The specialist or member who is highly talented in identification of the suspected disability may be more effective individually in making evaluation judgments than would be a group not containing a highly gifted member making similar judgments (Middlebrook, 1974).
5. An agenda assists in increasing the effectiveness of the assessment team. Tubbs and Moss (1974) describe the *Single Question Form* as an agenda which seems to best meet the spirit of PL 94-142. It is:
   a. What is the single question, of which the answer is the only thing the group must know to accomplish its purpose?
   b. What subquestions must be answered before we can answer the single question we have formulated?
   c. Do we have sufficient information to answer confidently the subquestions? (If yes, answer them; if not, continue below.)
   d. What are the most reasonable answers to the subquestions?
   e. Assuming that our answers to the subquestions are correct, what is the best solution to the problem? (Larson, 1969 p. 453).

Recommendations Related to Nondiscriminatory Testing

1. Examine placement recommendations or decisions for biases. Administrators of programs should examine their placement records. If the records indicate that individuals placed appear to be predominantly from certain cultural, ethnic, or sex groups, procedures for determining those placements should be examined carefully for cultural biases, and should be changed if biases are discovered.

   If the assessment procedures appear to follow the best guidelines but biases still appear to exist, perhaps other possible causes should be examined. Are the individuals responsible for actually acting on the placement recommendations arranging for placement of some children before others? Are some children recommended for returning to regular classes at a higher rate than others? If one can find no bases for bias in placement after careful search, one at least will have examined the situation thoroughly and be able to justify the existing placement procedures.

   Additionally, teachers may wish to examine the records for their own history of referring individuals for special class placement. If they had referred a statistically higher percentage of minority students, they perhaps should try to determine the basis upon which these recommendations were made.
2. Use test items that reflect the content of the curriculum. If the tests being used predict the probability that individual students will be able to succeed in school, one should determine if those tests are truly related to success. Although this is generally a re-

search issue to be addressed in other situations, one can grossly determine such a relationship by face validity.

For instance, if attempting to predict probability of success for an individual's completing a vocational education program, using a test which has been validated for predicting success of individuals in college preparatory programs, such a test probably is not useful for the first purpose. The style and content of test items for the two predictions probably would be quite different. If one suspects from examining a test that it is invalid for a particular purpose, one should check the manual to determine if the test has been used before for the current intended purpose. If data are unavailable, assume that the measure probably is not appropriate.

3. Results of an assessment battery should yield measures of both standardized and optional performance, and should indicate both competencies and lack of competencies. Many tests do not measure students' behavior under the best circumstances; this factor needs to be considered, and followed by any modifications needed. Results of the assessment should indicate not only what diagnosis is most appropriate, but also the individual strengths, discrepancies between standardized and optimal performance that may be important, areas of growth that should be targeted, etc.
4. Realize that a test samples only a small part of the child's behavioral repertoire. On any given behavior tested, one is measuring a small part of the child's total behavior. Thus, one is drawing inferences about the child's total condition on the basis of a small part, and this can lead to gross errors. The smaller the item sampling, the more likely there is item bias.
5. Motivational factors may affect a child's scores adversely, but inflating children's test performance through motivational factors is unlikely. Behaviors can be prompted, and the scores inflated, but motivation will be due to consequent events. A child's scores, however, may be much lower, because of motivation, than that of which she or he is capable. Therefore, motivation does need to be taken into account in interpreting test results adequately. If extrinsic reinforcers are used in a testing situation of standardized tests, however, they may well affect the results, since testing conditions were not standard.
6. Periodically evaluate your own empathy toward minority vs. majority groups. If you have either strongly positive or negative bias toward some group, you would do well to question your test results, or possibly check them with other testors.
7. If errors occur in placement decisions, they should be in the direction of least-restrictive placement. That is, if questionable or borderline cases arise, decisions should be made toward regular school class placement, followed by frequent monitoring to ensure that the placement is not affecting the child adversely. Additionally, children placed in special education classes should receive continuous monitoring, to be able to recognize sufficient improvement allowing for placement in a less restrictive class.
8. Make explicit criteria for placement in special classes, and evaluate how good these criteria are for making decisions.
9. One way to resolve the dilemma of minority children appearing to perform poorly on standardized tests is to set up new norms that would allow a child's raw score on the test to be translated to an appropriate IQ score with a mean of 100 and the same standard deviation as the majority form of the test. In this way, a child scoring 75 on majority norms would achieve an IQ of 100 on the norms for the minority group represented by that child *if* the performance represents normal functioning for a child in that sociocultural modality group. Using this system, the norms would be set up differently for each minority group in the same way that they currently are set up for children of different ages and sexes.

   This procedure, however, does not provide for analysis of defects in the test nor does it provide the opportunity for minority children to demonstrate excellence in abilities, information, and problem solving strategies that may be stressed within their sociocultural group but not in the majority's.

RECOMMENDATIONS FOR FUTURE RESEARCH

The above recommendations for practitioners are intended as usable guidelines which can be followed until

researchers approach the necessary questions systematically. After a solid data base has been developed on nondiscriminatory testing, the recommendations proposed here can be examined in light of the new data, and be accepted, modified, or rejected. Until those data are available, the authors hope the suggestions in this article will be helpful to those who are faced with immediate problems in educational assessment.

To facilitate development and implementation of the necessary research efforts, the authors have developed sampling questions that need to be addressed, and further suggest that the field would be enhanced greatly if these questions were addressed systematically by some group or agency. This research project could be a useful funding priority for the Bureau of Education for the Handicapped, Office of Education.

The research questions which seem essential to development of effective nondiscriminatory testing are:

1. What minimum common skills are necessary for survival in both majority and minority cultures, and how can these skills be measured? If these skills can be determined, can they beneficially be used as a predictor of the need for special education services?
2. How reliable are existing tests that have been developed or modified from older tests when they are used as suggested? If the tests are reliable, how valid are they for the intended purposes?
3. How are the norms of existing tests affected when these tests are restandardized on other minority cultures, or on the current majority culture? What validity do these tests have for predicting certain educational outcomes (such as failure in regular classes, but success with special services)?
4. What is the effect of using minority examiners on test results when controls for empathy and testing procedures are introduced? Do they affect the interpretation of results? Of overall norms?
5. What performance standards are most appropriate with various tests, including criterion-referenced measures?
6. Are there theoretically-oriented differences between cultures that affect test results? (For example, do all cultures progress through the same development sequence? Would a Piagetian developmental test be equally appropriate for all cultures?)
7. What is the efficiency of PL 94-142 on nondiscriminatory placement and testing?

These questions represent a small part of what must develop into a major effort before nondiscriminatory testing becomes a reality. The authors hope that this effort is eminent.

REFERENCES

Alley, G. R. *Methods of assessing the performances of exceptional children who represent minority groups.* Paper presented at Regional Problems of Exceptional Children Meeting, Regional USEO and GAC, Kansas City, MO, June 1976.

Anderson, J.E., & Alley, G.R. *A comparative study between educably mentally retarded and nonhandicapped children on visual-motor tasks and metamemory self-report performances.* Manuscript submitted for publication, 1977.

Balthazar, E.E. *Balthazar scales of adaptive behavior, Part 1: Handbook for the professional supervisor.* Champaign, IL: Research Press Co., 1971, (a)

Balthazar, E.E. *Balthazar scales of adaptive behavior, Part 2: Handbook for the rater technician.* Champaign, IL: Research Press Co., 1971. (b)

Balthazar, E.E. *Balthazar scales of adaptive behavior, Part 3: The scales of functional independence.* Champaign, IL: Research Press Co., 1971. (c)

Balthazar, E.E. *Balthazar scales of adaptive behavior, Part 4: Workshop and training manual.* Champaign, IL: Research Press Co., 1971. (d)

Balthazar, E.E. *Balthazar scales of adaptive behavior II: Scales of social* adaptation. Palo Alto, CA: Consulting Psychologists Press, 1973.

Budoff, M., & Hutton, L. The development of a learning potential measure based on Raven's Progressive Matrices. *Studies in Learning Potential,* 1972, *1* (18).

Cain, L.F., Levine, S., & Elzey, F.F. *Manual for the Cain-Levine social competency scale.* Palo Alto, CA: Consulting Psychologists Press, 1963.

Davis, J. *Group performance.* Reading, MA: Addison-Wesley, 1969.

Doll, E.A. *Vineland social maturity scale—Manual of directions.* Minneapolis: American Guidance Service, Inc., 1947.

Flavell, J. H. Developmental studies of mediated memory. In L. P. Lipsitt & H. Reese (Eds.), *Advances in child development and behavior* (Vol. 5). New York: Academic Press, 1970.

Foster, R. W. *Camelot behavioral checklist manual.* Parsons, KS: Camelot Behavioral Systems, 1974.

Foster, R. W., & Barnes, K. *Range of motion checklist.* Unpublished assessment instrument, 1977. (Available from Ray Foster Project LGARN, Kansas Neurological Institute, 3107 West 21st Street, Topeka, KS 66604.)

Fulton, R.T., & Spradlin, J.E. Operant audiometry with severely retarded children. *Audiology,* 1971, *10,* 203-211.

Garcia, R.L. *Unique characteristics of exceptional bilingual students.* Paper presented at regional meeting of USOE and GAC, Kansas City, MO, June 9, 1976.

Janis, I. Groupthink. *Psychology Today,* 1971, *5.* 43-46, 74-76.

Larson, C. Forms of analysis and small group problem solving. *Speech Monographs,* 1969, *36,* 452-455.

Leland, H., Nihira, K., Foster, R., Shellhaas, M., & Kagin, E. Conference on Measurement of Adaptive Behavior: II. Parsons, KS: Parsons State Hospital and Training Center, 1966. (Monograph)

Lerner, J. W. *Children with learning disabilities: Theories, diagnosis, teaching strategies* (2nd ed.). Boston: Houghton Mifflin Co., 1976.

Mercer, J.R. *Labeling the mentally retarded.* Berkeley, CA: University of California, Press, 1973.

Messick, S., & Anderson, S. 2. Educational testing, individual development and social responsibility. In R.W. Tyler & R.M. Wolf, *Crucial issues in testing.* Berkeley, CA: McCutchan Publishing, 1974, pp. 21-34.

Middlebrook, P.N. *Social psychology and modern life.* New York: Alfred A. Knopf, 1974.

Nihira, K., Foster, R., Shellhaas, M., & Leland, H. *AAMD adaptive behavior scale manual.* Washington, DC: American Association of Mental Deficiency, 1974.

Platt, J.S. *The effect of the modified Raven's Progressive Matrices Learning Potential Coaching Procedure on Raven's posttest scores and their correlation value with productive variables of learning disabilities.* Unpublished doctoral dissertation, University of Kansas, 1976.

Sailor, W., & Horner, R.D. Educational assessment strategies for the severely handicapped. In N.G. Haring & L.J. Brown (Eds.), *Teaching the severely handicapped* (Vol. 1). New York: Grune & Stratton, 1976.

Sailor, W., & Mix, F. J. *The TARC assessment system* (user's manual). Lawrence, KS: H & H Enterprises, 1975.

San Antonio Independent School District v. Rodriquez. 410 U.S., 93 S. Ct. 1278, 36 L. Ed. 2d 16 (1973).

Shearer, D., Billingsley, J., Frohman, A., Hilliard, J., Johnson, F., & Shearer, M. *The Portage guide to early education: Instructions and checklist* (Experimental ed.). Portage, WI: Cooperative Educational Service Agency No. 12, 1970.

Somerton, E., & Turner, K. *Pennsylvania training model: Individual assessment guide.* King of Prussia, PA: Regional Resources Center of Eastern Pennsylvania for Special Education, 1975.

Spellman, C.R., & DeBriere, T. *Research and development of subjective visual acuity assessment procedures of severely handicapped persons* (BEH Grant No. G0076 02 592). Parsons, KS: Parsons State Hospital and Training Center, 1976.

Stacker, J. Se Habla Espanol. *American Education,* May, 1967, 17-18.

Swanson, J.E. *Learning potential as a predictor of behavioral changes in learning disabled elementary students.* Unpublished Master's thesis, University of Kansas, 1976.

Thorndike, R.L. 3. Concepts of culture-fairness. In R.W. Tyler & R.M. Wolf, *Crucial issues in testing.* Berkeley, CA: McCutchan Publishing, 1974, pp. 35-45.

Tubbs, S.L., & Moss, S. *Human communication: An interpersonal perspective.* New York: Random House, 1974.

Wagner, H. Attitudes toward and of disadvantaged students. *Adolescence, 7* (28), 1972, 435-446.

Wechsler, D. Manual for the Wechsler Adult Intelligence Scale. New York: Psychological Corp., 1955.

Williams, R.I. 1. Black pride, academic relevance and individual achievement. In R.W. Tyler & R.M. Wolf, *Crucial issues in testing.* Berkeley, CA: McCutchan Publishing, 1974, pp. 13-20.

CLASSROOM FORUM

Edited by Debby Gilbert
Curriculum Coordinator
Learning Disabilities Program
DeKalb County, Georgia

I have found that students with specific learning disabilities sometimes enter high school lacking certain academic skills and habits which would enable them to function well in high school. The performance of such students is often marginal or unsatisfactory. How can I help?

A study skills program can be introduced with entering eighth grade students, designed so that students with learning deficits will gain many compensatory learning skills. Other students desiring to improve their study habits and skills also would benefit from the course described below.

The students would meet as a group for one hour each school day, and would receive a grade for the course. The course would be repeated each term for another group of 15 or 20 students. The course would begin with a reading improvement unit, and then cover vocabulary building and enrichment, notetaking, memory improvement, logical thinking, test-taking techniques, spelling improvement, establishing good study habits, and other units as time permits.

Students would keep charts indicating their reading comprehension scores and progress. The *Nelson Reading Tests,* Forms A and B, (New York: Houghton Mifflin Company, 1962) could be used as a pre-test and a post-test. Informal reading tests measuring progress in vocabulary and comprehension would supplement the Nelson tests.

The vocabulary unit might consist of graded sight vocabulary words. Idioms, context clues, using the dictionary, affixes, synonyms, antonyms, homonyms, definitions, and the use of the thesaurus are other suggested inclusions, along with ways to build and enrich vocabu-

lary. Students should be encouraged to become interested in the etymology of words, to learn specialized words, and to collect new words.

In the test-taking unit, the emphasis should be against cramming, toward spacing study sessions, and breaking the material into equalized segments for studying. Good study habits may be encouraged by giving the students a checklist of good study habits and having them score themselves. The SQ3R study method may be taught for use with social studies and other subjects. Other units, as time permits, may include using the library, preparing and giving oral reports and talks, and extending study skills into the content areas.

This "How to Learn" course can improve the abilities of students with learning disabilities, and other students can become more efficient in their study methods, resulting in improved grades. For all students taking part, they will be able to work at levels more nearly approximating their capacities.

We thank Mrs. Anne Perry, Learning Disabilities Resource Teacher, Lakeside High School, DeKalb County, Georgia, for providing information on this innovative project. She developed the above-described program, which has resulted in a high rate of success with learning disabled and other students at the eighth grade level.

I find that the six-year-olds who come to my diagnostic class often have not yet learned self-help tasks like tying their shoelaces and hanging up their coats and sweaters. Rather than being frustrated by following them around doing these things for them, what can I do to help them learn these skills as quickly as possible?

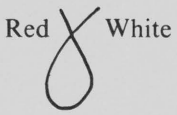
In the case of special children, their lack in basic self-care skills upon entering school is not always due to lack of instruction at home or in preschool programs. Most often it is because these children do not learn to perform these tasks in the conventional manner or at the same rate as the majority of children. They fail to learn from repeated demonstration by concerned parents or from the usual group instruction given in preschool programs or even from continual daily observation of other children. Rather than your attempting to grow an extra pair of hands and continuing to help these students, try these sequences:

Hanging up coats and sweaters

1. Child observes instructor hanging up coat; instructor verbalizes the steps involved while hanging up the coat.
*2. Child lays his or her coat flat on the floor with sleeves extended, front side up.
*3. Child places a hanger into the sleeves of the coat on the floor.
*4. Child folds each side of the coat front toward the middle of the hanger.
5. Child then hangs the coat on the coat rack.
6. Instructor holds the hanger as the child places the coat appropriately on the hanger, sleeve by sleeve; the child then hangs the coat up on the rack.
7. Child holds own hanger, places each sleeve of the coat on the appropriate side of the hanger, then hangs it up on coat rack.

Tying shoelaces

1. Half the demonstration lace should be white and the other half red or other contrasting color to simplify distinguishing right from left.
2. Instructor shows child how to cross the right side over the left, verbalizing: "The worm is going to crawl over his tail."



3. Instructor shows child how to bring the right string over the left and up through the loop created, ver-

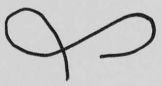
*May need verbalization from the instructor until the steps are automatic for the child.

balizing: "The worm peeks down into the hole to see Mr. Rabbit."

4. Instructor shows child how to pull the string down into the loop, verbalizing: "Now pull the worm into the hole and hold his tail tight."

Making a bow (as in tying shoelaces)

1. Instructor shows child how to make the two loops of the bow, verbalizing: "Make two rabbit ears like this, crossing the red ear over the white ear."



2. Instructor shows child how to push the right loop through the left, verbalizing: "Poke the red ear through the hole."
3. Instructor shows child how to pull the bow tight, verbalizing: "Hold both ears so they don't get lost, pull slowly, and pull tight."

Additionally, children may need to work on a shoe frame (one you have made or one available commercially), lacing cards, or their own shoes with regular laces after lots of practice.

We thank Mrs. Nancy Halpern, Learning Disabilities Teacher, Oak Grove Elementary School, DeKalb County, Georgia, for contributing her learning sequences to this column.