Identification of the Academically Gifted

LYNN H. FOX Johns Hopkins University

ABSTRACT: Scientists and public officials have sought to define giftedness in a way that will ensure that outstanding academic potential in schoolchildren is encouraged. Various criteria for identifying the academically gifted, such as scores on general intelligence and creativity tests, teacher recommendations, and scores on standardized achievement tests, have been used. Fox points out their limitations and recommends an identification process developed by Stanley, which equates precocity with academic talent by focusing on children with exceptionally high performances on advanced tests of specific subject matter. Fox discusses various programs that begin with this process and then supplement it with further diagnostic testing, clinical methods, and evaluation of students' products. She notes that use of these procedures with disadvantaged populations has identified more academically gifted students than other procedures have found. —The Editors

Although Galton (1869) was the first to systematically address the concept of genius in the psychological literature, a history of psychological testing records the use of tests to discover talented individuals in China as early as 2200 BCE (DuBois, 1970). While Guy M. Whipple has been credited with first using the term gifted to designate children with superior intellectual ability (Passow, 1981), Lewis M. Terman (1925) initiated the first major research study in which giftedness or intellectual talent in children was operationalized in terms of performance on intelligence tests. Terman sought students who scored in the top one percent on the 1916 version of the Stanford-Binet Intelligence Scale.

Giftedness: Concept and Definition

Most conceptual definitions of giftedness include some reference to intelligence, without attempting to define the term. This is perhaps not surprising in light of the general lack of consensus about the meaning of the term beyond "that which an intelligence test measures" (Thorndike, 1921). Wechsler (1975) suggests that intelligence tests measure "the capacity of an individual to understand the world about him and his resourcefulness to cope with its challenges" (p. 139). It may be adaptability to the school environment in particular that characterizes students who have high scores on intelligence tests, or as Leta S. Hollingworth is quoted as saying in 1931, "By a gifted child, we mean one who is far more educable than the generality of children are" (cited in Pritchard, 1951, p. 49).

Renzulli (1978) noted that definitions of gifted differed in terms of the number of performance areas that are specified and/or in the degree of excellence that must be exhibited. For example, the operational definition used by Terman would be considered highly restrictive. Several, much broader definitions have been proposed, such as that the talented and gifted child is "one who shows consistently remarkable performance in any worthwhile line of endeavor" (Witty, 1958, p. 52).

Definitions also differ on whether or not evidence of potential alone is a sufficient condition for giftedness. According to some definitions, a child who exhibited potential but was performing poorly in school or on achievement tests still could be considered gifted. Other definitions, however, require that the child demonstrate high levels of skills or knowledge in addition to potential. For example, Fliegler and Bish (1959) define gifted students as those who possess both potential and intellectual talent in children was operationalized in terms of performance on intelligence tests. Terman sought students who scored in the top one percent on the 1916 version of the Stanford-Binet Intelligence Scale.

The Study of Mathematically Precocious Youth (SMPY), initiated in 1971 by Julian C. Stanley at The Johns Hopkins University, introduced a new perspective on the definition of giftedness by identifying students who were talented in specific academic areas (Stanley, Keating, & Fox, 1974).
SMPY sought youths who reasoned extremely well in mathematics at an early age, as evidenced by high scores on difficult tests (those intended for much older students), and so equated precocity, as measured by test performance, with academic talent. Stanley (1976c), however, noted that scoring well beyond one's chronological age on a difficult test implies not just earlier development, but also higher levels of ability that "presage long-range, lasting differences in ultimate ability" (p. 6). Thus, a student might be identified as academically gifted in one or more specific areas without necessarily exhibiting general intellectual superiority overall. The numbers of students identified would be a function of the number of talent areas considered and the extent of overlap among them.

Although the terms academically talented and gifted are often used interchangeably, Renzulli (1978) has argued that giftedness should be defined in terms of the interaction of three clusters of traits: above-average general ability, high levels of task commitment, and high levels of creativity. Accordingly, "gifted and talented children are those possessing or capable of developing this composite set of traits and applying them to any potentially valuable area of human performance" (Renzulli, 1978, p. 261). The rationale for this concept is based on the cumulative research on creative and productive adults (Albert, 1975; Cox, 1926; MacKinnon, 1965; McCurdy, 1957; Roe, 1952). Presumably, this definition would identify far fewer students than the ones proposed by Stanley et al. (1974) or Witty (1958).

Sternberg (1981) suggests that giftedness should be described in terms of an information-processing theory of intelligence rather than a psychometric one. According to his theory, there are three broad components of giftedness: metacomponents, such as the higher-order processes used in problem solving (e.g., selection of strategies for problem solving); performance components, the processes used in problem solving, such as inference; and acquisition, retention, and transfer components, the skills used in learning, storing, and applying information. Gifted individuals are those who are capable of "high quality and quantity of interaction among the various kinds of components in the system . . . they are more sensitive to the feedback that the various components can provide" (p. 91).

The most extreme definition has been proposed by Feldman (1979), who argues that "all children are gifted" (although not all in the academic and intellectual areas). He says, "farfetched though it may sound, I think it not at all implausible that a broadened view of giftedness would reveal that every child is gifted in some socially valued way." This is radically different from the "sociopsychological approach" proposed by Newland (1976), in which a certain percentage of students would be identified on the basis of certain behavioral characteristics that correspond to those needed, according to a hypothetical assessment, for society's future leadership.

**Legal Definitions**

Prior to 1971, only four states (California, Illinois, Nebraska, and Pennsylvania) had laws or regulations in which giftedness was defined. Interest in the academically gifted had been aroused on a national level in the late 1950s as a reaction to the launching of Sputnik by the Soviet Union. Educational efforts, however, were focused more on curriculum development in mathematics and science and on teacher training than on identification of highly able students. A report to Congress in 1971 by the Commissioner of Education sparked new interest in the identification of talented students (Marland, 1972). This document called gifted children a neglected national resource and also questioned the reliance on intelligence test scores as the sole criteria for determining talent. For the first time, a definition of gifted and talented children was proposed at the national level (Pub. L. 91-230, § 806), as follows:

Gifted and talented children are those identified by professionally qualified persons who by virtue of outstanding abilities, are capable of high performance. These are children who require differentiated educational programs and/or services beyond those normally provided by the regular school programs in order to realize their contribution to self and society.

Children capable of high performance include those with demonstrated achievement and/or potential ability in any of the following areas, singly or in combinations: (1) general intellectual ability, (2) specific academic aptitude, (3) creative or productive thinking, (4) leadership ability, (5) visual and performing arts, (6) psychomotor ability.

Subsequent federal legislation dropped the sixth category of psychomotor ability. By 1978, 42 states either had developed definitions for giftedness in formalized law or regulation or had at least formulated working guidelines modeled along the lines of the federal definition (Karnes & Collins,
General intellectual ability was specified in all of the 42 states. In 36 states, the definitions included separate categories for specific academic aptitude and creative thinking.

**Operational Definitions**

Three states define intellectual giftedness as a score that is two standard deviations above the mean on an individual test of intelligence. In most states the definitions are less specific and refer to students who score in some upper range (2% to 10%, or one-and-a-third to two standard deviations above the mean) on measures of ability (Karnes & Collins, 1978). The federal definition requires only that students be identified by experts.

Although experts typically recommend the use of individual tests of intelligence such as the Wechsler Intelligence Scale for Children-Revised (WISC-R) or the Stanford-Binet, in practice most local school systems rely on a combination of teacher recommendations and scores on group tests of intelligence and achievement. In a 1971 survey of 204 experts in education, 49% felt that the use of individual intelligence tests was essential, whereas group tests of intelligence, group achievement tests, and teacher judgments were viewed as essential by only 12%, 18%, and 19%, respectively. In a survey of actual school practice in school systems in Illinois, however, only 23% reported using individual tests of intelligence, whereas group tests of intelligence and achievement were used by 87%, and teacher nominations were used by 93% of the school systems (Marland, 1972).

That schools generally rely on group tests rather than on individual tests for identification is not surprising given the high cost in time and money of administering individual tests and the general lack of federal, state, or local monies for education of the gifted and talented. Indeed, prior to 1972 most schools did not even have any procedure for identification of these students. In a School Staffing Survey (Marland, 1972), 57.5% of the schools surveyed reported that they had no gifted students enrolled.

An operational definition for specific academic aptitude proposed by Stanley (1979) begins with an initial screening on an in-grade achievement test. Students who score in the upper 3% qualify for further testing on a more difficult measure of specific aptitude. The criterion score on the more difficult test necessary to qualify for the SMPY programs has varied depending on the test used, the nature of the program, and the age of the student, but as a general rule, seventh graders who scored at least 430 on the Scholastic Aptitude Test—Verbal subscale (SAT-V) or 500 on the Scholastic Aptitude Test—Quantitative subscale (SAT-M) have been considered academically talented. Although the talent searches at Hopkins have relied heavily on the use of the SAT, Stanley (1979) finds that the Differential Aptitude Test, the Preliminary Scholastic Aptitude Test, and the American College Testing Program can be used in place of the SAT. Final placement in programs involves the use of additional information about the student, such as measures of interest and achievement on tests related to the content of specific courses.

**Validity**

Evidence of the long-term predictive validity of individual intelligence test scores for identifying the academically gifted comes largely from Terman's longitudinal study of children with IQ scores of 140 or higher on the Stanford-Binet. As adults, the group out performed other comparison groups on the Concept Mastery Test, a difficult test of verbal reasoning. The proportion of the Terman group members who graduated from college, earned advanced graduate degrees, and pursued professional careers was considerably larger than that of the general population (Oden, 1968). The numbers who published articles and books, held patents, or earned awards and honors were also impressive.

Of course, there were some subjects in the Terman study who were not especially successful. It could be argued that academic potential assessed in childhood is a necessary but not sufficient condition for eminence in life. Indeed, studies of eminent adults suggest that while some minimal level of academic talent may be essential, other factors not assessed by intelligence tests are also necessary to enable individuals to become highly productive and creative (Albert, 1975; MacKinnon, 1965). It could also be argued that individuals who show potential in childhood may not develop that potential because their talents were not nurtured in the home or school. Indeed, it was found that some educational experiences and home factors differentiated between the most and the least successful men in the Terman study (Oden, 1968).

Even though scores on group tests often correlate
highly with scores on individual tests of intelligence, group tests often do not have enough ceiling to differentiate well among the most able students. For example, Pegnato and Birch (1959) reported that for junior high school students, a cutoff score of 130 on a group test (Otis Quick-Scoring Mental Ability Test) identified only 22% of those who scored 136 or higher on the Stanford-Binet, whereas a cutoff of 115 correctly identified 92%. In another study of 332 gifted students who scored 130 or higher on the Revised Stanford-Binet scale, only half scored 130 or higher on a group test of intelligence (Martinson & Lessinger, 1960). Thus, it may be wise to use scores of 120 or even 115 on group tests for purposes of screening, but not for decisions about program placement.

If the purpose of identifying academically talented students is to provide them with appropriately difficult and challenging educational programs, there are then several arguments, based on content validity, against the use of intelligence tests. Measures of global intelligence do not provide sufficient information about patterns of specific abilities and levels of achievement for program placement. For example, two students who score 130 on the Stanford-Binet may have different patterns of ability and different levels of achievement in specific school subjects. One student may be very gifted in mathematics and have mastered content well beyond grade level, while the other student is only moderately good in mathematics, but extremely talented in language arts. Thus, for diagnostic and prescriptive purposes, batteries of aptitude and achievement tests are more useful than intelligence tests. It has also been argued that the contents of such tests are biased against students in some ethnic groups and do not measure many important cognitive abilities such as divergent thinking skills.

Considerable evidence in support of the validity of difficult tests of specific aptitude and achievement for educational placement of gifted students has been amassed by SMPY and its parallel project, the Program for Verbally Gifted Youth (PVGY). For example, scores on the SAT-M in grade 7 or 8 are clearly related to success in accelerated mathematics classes (Fox, 1974; George & Denham, 1976; Stanley, 1980), and even predict the degree of acceleration students will accomplish in a short period of instructional time (Bartkovich & Mezynski, 1981). In the verbal area, scores on the SAT-V in combination with the Test of Standard Written English (TSWE) in grade 7 can be used to select students who can master a college level German course in fewer hours but at as high or higher a level of mastery as college students in a regular class (Durden, 1980; McClain & Durden, 1980). High scores on the SAT-M or SAT-V have also been used successfully to select seventh and eighth graders for college courses on a part-time basis (Solano & George, 1976) and used as criteria for early admission to college, even radical acceleration by four or five years (Stanley, 1976a).

Several school systems have replicated the SMPY model using tests like the SAT, PSAT, or SCAT to select students for accelerated mathematics classes (Solano, 1979; Fox, Brody, & Tobin, Note 1). Efforts to replicate the verbal courses have also been successful (Van Tassel-Baska, Note 2). A large-scale replication of the Johns Hopkins talent search and special classes program was conducted in the summer of 1981 at Duke University (Sawyer, Note 3).

There is little research on the long-term power of tests of specific aptitude and achievement to predict life accomplishments. The ongoing SMPY longitudinal study has found, however, that scores on difficult tests of mathematical and verbal reasoning (SAT-M and SAT-V) taken in grade 7 and 8 correlate with aptitude scores and several measures of achievement in high school (Benbow, 1981). Stanley (1976c) says, "for appropriate criteria, validity does not drop at the upper part of the score range of a test that is difficult enough for the persons tested" (p. 5).

There are little data with which to compare the predictive validity of teacher judgments versus standardized tests. Some studies have contrasted teacher nominations with the criterion of test performance. In the Terman study, teachers were asked to nominate the first, second, and third brightest children in their class and also to name the youngest. Nominations of the youngest child in the class produced more students who met the criterion of 140 or higher on the Stanford–Binet than did nominations of the brightest (Terman, 1925). It is important to recall, however, that the majority of the Terman sample was nominated in one of the categories by teachers. Terman estimated that, in some schools, 10% to 25% of the students who would have a score of 140 or higher were not nominated by teachers. There is no way to know whether those students achieved as well in adulthood as the ones who were nominated by teachers.

Gear (1976) reviewed five studies in which teacher nominations of students were compared.
with a criterion of student scores two standard deviations or more higher than the mean of an individual test of intelligence, such as the WISC-R or the Stanford-Binet. In all five studies, teachers correctly identified less than half of the students, and in one study, they only identified 10% (Jacobs, 1971). Teacher judgments were also inefficient in that they identified 10% or more of the nongifted students as gifted. Borland (1978) found teacher ratings on a checklist to be about as efficient as scores on group tests of intelligence.

Stanley (1976b) found that difficult tests of mathematics taken in grades 7 or 8 were better predictors of performance in a mathematics contest in grade 11 than judgments of the 11th-grade mathematics teachers. Ten of the 51 students who entered a contest in grade 11 were nominated because of high test scores in grade 8. The 10 ranked 1, 2, 3, 5.5, 7, 8, 12, 16.5, 19, and 23.5.

Since the vast majority of teachers has never been required to take courses in the area of education for the gifted, it seems logical to assume that they are not knowledgeable about the characteristics of such students and may nominate dutiful students rather than those who are brilliant but bored. Gear (1978) studied the effect of specific training about the characteristics of gifted children on teacher effectiveness and efficiency. Training increased effectiveness (the number of children nominated by the teachers who met the test criterion) but not efficiency (the proportions of truly gifted among the total groups nominated by the teachers).

Group Differences

Although academically talented students can be found in all ethnic groups when standardized tests are used for selection, there are differences in the proportions identified in some ethnic groups as opposed to others. Whether or not group differences in performance on intelligence tests reflect real differences in intellectual potential or merely reflect cultural differences has been widely debated (e.g., Garcia, this issue; Gordon, this issue). Most educators in the gifted education field would argue for the use of alternative strategies to identify gifted students in populations (such as black, Hispanic, and American Indian groups) that are considered to be educationally disadvantaged or culturally different. At present, there is no evidence on the effectiveness of the alternatives.

One of the methods suggested for identifying gifted students among the culturally different or disadvantaged is the use of only selected subscales of intelligence tests—usually the nonverbal and performance ones. An Abbreviated Binet for the Disadvantaged (ABDA) was devised by Bruch (1971) for disadvantaged black children, and the System of Multicultural Pluralistic Assessment (SOMPA) was devised by Mercer and Lewis (1978) to adjust test scores in relation to a child's sociocultural group. Meeker and Meeker (1973) have used the Structure-of-Intelect model to specify patterns of strengths and weaknesses on the Stanford-Binet for black and Chicano students.

Le Rose (1978) has argued for a simple quota system approach to the problem whereby the top 3% to 5% of every subgroup would be screened for giftedness regardless of the difference in performance between groups on the specific test used. Others have argued for totally abandoning the use of standardized intelligence and achievement tests. For example, Torrance (1971, 1977) has recommended using the Torrance Tests of Creative Thinking to identify the disadvantaged gifted instead of using traditional intelligence measures. Self-report inventories have also been suggested, such as the Relevant Aspects of Potential (Grant & Renzulli, 1975) and the Alpha Biographical Inventory (Fraiser, 1979; Sato, 1975). Checklists for parents and teachers and peer nominations have also been advocated (Sato, 1975).

Although mean scores for boys and girls do not differ on measures of global intelligence, there do appear to be some gender-based differences in performance on tests of some specific abilities and achievement (Maccoby & Jacklin, 1974). Among the gifted, the most notable sex difference is in the proportion of boys who score higher than girls on the SAT–M at grades 7 or 8 (Fox, 1977; Fox & Cohn, 1980). No differences are found on the SAT–V for these same students. Although one might speculate about the social or biological basis for sex differences in achievement and aptitude, there is no conclusive evidence to support or refute either explanation.

The Creativity Controversy

Passow (1981) noted an increasing emphasis on creativity (defined in many ways) as either a component of academic and intellectual giftedness or as a type of giftedness itself. Guilford (1950, 1967) was the first to point out that some types of think-
ing involved in creative problem solving are not assessed by typical achievement and aptitude tests. Eminent adults were not always obviously precocious in childhood (Cox, 1926) and may not necessarily score very high as adults on tests of intelligence (Albert, 1975; MacKinnon, 1965). The work of Getzels and Jackson (1958) and Torrance (1965, 1977) suggests that some children who score only moderately high on intelligence tests, but high on creativity measures, are capable of high levels of achievement. On some tests of creativity, there are few, if any, differences among cultural groups (Torrance, 1971).

Whether or not it is possible to predict adult creativity from behaviors measured in childhood or adolescence is unclear (Crockenberg, 1972). An ongoing longitudinal assessment of the Torrance Tests of Creative Thinking (Torrance, 1977) may soon provide some data of interest. It has been argued also that creativity cannot be demonstrated before there is mastery of a discipline, and the ability to master a discipline is better measured by tests of aptitude and achievement than by tests of creativity.

Some of the controversy about the concept of creativity surrounds the question of the extent to which it is a personality trait, a specific cognitive ability, or a type of problem-solving strategy that might be learned (Michael, 1977). If it is the latter, there is less need to search for those who know the process than to teach the process to everyone.

**Implications and Implementations**

On the basis of existing research evidence, the most defensible approach for identifying the academically talented would appear to be the use of a variety of psychometric and nonpsychometric measures for initial screening. The information and methods used to screen a school population should include: evidence of high academic potential as assessed by standardized intelligence and aptitude tests; evidence of specific academic achievement as measured by standardized tests and/or informal measures such as teachers’ or experts’ judgments about science projects, essays, or poems; and recommendations of teachers, counselors, reading specialists, or other educators who have observed a student’s performance in a learning environment. Peer nominations, parent nominations, and self-nominations as well as performance on tests of creativity might also be used in the initial screening stage.

Decisions about the placement of students in special programs for the gifted or for the individualization of instruction within the regular class should be made by a qualified expert, or team of experts, on the basis of diagnostic testing, clinical methods such as interviews, and evaluation of products such as science fair entries, essays, poems, or short stories. Diagnostic testing for program placement should utilize the principle of difficult tests to discover potential as developed by Stanley et al. (1974). Any student who scores in the upper 5% or even 10% on an in-grade test should be tested on increasingly advanced levels of the test to determine her or his appropriate instructional level. The educational needs of a child can best be determined when there is a great deal of information available about both patterns and levels of abilities, as well as interests. Although there is no single perfect test of academic talent, test performance is still an efficient and effective method for identifying the needs of the majority of talented students. Efforts to find valid indices of talent in disadvantaged or culturally different populations must, however, be continued. Two model programs in which multiple methods are employed for “culture-fair” screening are described in the following paragraphs.

The Gifted and Talented Education (GATE) project in the Baltimore City Public Schools was one of the first efforts to implement the multiple screening and identification procedures derived from the federal definition of giftedness. The program gives special attention to minority and disadvantaged populations. Each participating school has a screening committee consisting of five or more school personnel. Systematic procedures for screening in each of six categories (intellectually gifted, academic aptitude, visual and performing arts, creative and productive theory, leadership, and psychomotor) at the elementary and secondary levels are provided for the committees (Cooke, Note 4; Cooke, Coplen, & Knox, Note 5). Behavioral checklists have been developed for all six categories of giftedness at both the elementary and secondary level and can be used by parents, teachers, and other appropriate persons to nominate students. Standardized tests, auditions, evaluations of student projects, and interviews also are used in the screening process. Final decisions for student placements are made on the basis of a composite profile analysis and a committee interview with each student, and if needed, with the parents.

An example of the screening procedures in the
GATE project for intellectual giftedness at the secondary level follows. Students can be referred to the screening committee on the basis of (a) scoring in the top 6% on a group test such as the Raven's Standard Progressive Matrices; (b) checklists completed by a teacher; (c) scoring two or more years above grade level on an achievement test such as the Iowa Test of Basic Skills; and (d) enrollment in enrichment classes regardless of test scores. Students who meet one of the above criteria are further screened on the Raven's Advanced Progressive Matrices and the WISC-R. Typically, scores of 125 on the verbal, performance, or full scale are required for all except the disadvantaged to obtain placement in the general intellectual category. For the disadvantaged, scaled scores of 14 are required on form subtests of the WISC-R (similarities, comprehension, block design, and picture arrangement).

The Flint, Michigan, school system is another that has attempted, with the help of the Educational Testing Service, to develop multiple screening methods to broaden the definition of giftedness beyond a single score on an intelligence test in an effort to increase minority participation in special programs. Students can initially qualify in one of three ways: on the basis of scoring two or more years above grade level on the reading and mathematics portions of the SRA Achievement tests; on the basis of teacher nomination in one of six areas of giftedness (creativity, leadership and social self-awareness, intellectual ability, learning potential, motivation, and exceptional academic achievement); or by scoring at or above the 90th percentile in their school and grade on a questionnaire of behavioral indicators completed by their parents. Following this screening, a detailed analysis of the child is compiled by the screening committee using behavioral checklists and tests as necessary.

The preliminary reports on the Flint project are encouraging. In 1977, before the project began, 62% of the students in the Flint schools were black, but blacks comprised only 26% of those identified as gifted. The percentage of black students increased to 44% in the pilot project. When children identified by the new procedures were compared with those identified by the old methods, the differences in performance in the programs were statistically insignificant (Lewis, 1980).

Any problem of classification or selection requires that a value judgment be made about the problems of false positives and false negatives. In the case of the academically talented, it might be wise to risk identifying students as gifted when they are not and give them the opportunities for special or individualized educational programs rather than to err by overlooking many talented students who are bored, frustrated, and unchallenged in their classes. Tannenbaum and Passow (cited in Passow, 1981) have suggested that perhaps students should be given the opportunity for some enriching educational experiences and then further evaluated on the basis of their performance in these learning environments.

One might envision a time in the future when diagnostic assessment and prescriptive instruction are developed for all children in keeping with their unique learning styles and instructional needs. Then there would be no need to try to predict whether or not a student is academically gifted, as each student would automatically be continually directed to more difficult and challenging learning experiences. Until that time, we must be satisfied with what realistically can be done to locate and challenge the academically talented children in today's classrooms.

REFERENCE NOTES


REFERENCES

Bruch, C. B. Modification of procedures for identification of
R. T. Zaffran (Eds.), *New voices in counseling the gifted.* Dubuque, Ia.: Kendall/Hunt, 1979.


Thorndike, E. L. Intelligence and its measurement: A symposium. *Journal of Educational Psychology,* 1921, 12, 124–133.


Torrance, E. P. Are the Torrance tests of creative thinking biased against or in favor of "disadvantaged groups?" *Gifted Child Quarterly,* 1971, 15, 75–80.

