Chapter 5
Equitable Assessment of Gifted Students Using the Naglieri General Ability Tests: Verbal, Nonverbal, and Quantitative

Jack A. Naglieri
University of Virginia, USA

Kimberly A. Lansdowne
Arizona State University, USA

ABSTRACT

Even though it is well documented that Black, Hispanic, Native American students have been denied access to gifted education for decades, injustice continues. The authors present research showing that traditional IQ tests with their verbal and quantitative questions contribute to under-representation because they yield large differences for students of color. Some (e.g., NNAT), but not all, nonverbal tests help but verbal and quantitative content is omitted. The authors suggest that students of color who are intellectually capable (gifted) but perhaps not talented (knowledgeable) could be more equitably evaluated if the verbal comprehension of instructions and verbal and quantitative knowledge were taken out of the tests used for identification. Research evidence is provided which shows that the Naglieri General Ability Tests: Verbal, Nonverbal, and Quantitative, which have nonverbal directions, do not demand knowledge acquired at school and do not require verbal response yield small differences by gender, race, ethnicity, and parental education.

INTRODUCTION

This book chapter provides a critical analysis of the procedures used to identify gifted and talented students. We describe gifted students as those with high general ability (regardless of their academic skills) and talented students as those who have high achievement (advanced academics). This distinction has

DOI: 10.4018/978-1-7998-8153-7.ch005

considerable implications for the types of tests used in the identification process. That is, five out of the six most widely used tests to identify gifted and talented students have characteristics that lead to under-representation of black, Hispanic, Native American and ELL students. We propose that test directions, content, and student verbal responses can serve as an obstacle to identification of gifted students. Our solution is an innovative method for identifying giftedness more equitably using a new trio of general ability tests that have verbal, quantitative and nonverbal content without (a) test questions that require knowledge, (b) the use of a specific language, (c) verbal instructions that demand comprehension of verbal concepts, and (d) the requirement that the student orally articulate their answers. Recent research (Selvamenan, et al., submitted for publication) using a large representative sample of students in grades K-12 (N = 8,105) demonstrates that these measures of general ability are more equitable across race, ethnicity and parental educational levels.

The objectives of this chapter, therefore, are to examine:

- how many gifted students of color could have been identified as gifted but were not;
- the relationship between test content and equity defined as rates of identification and mean score differences;
- how traditional verbal, nonverbal and quantitative approach can be modified so that the tests yield scores that are equitable;
- the role normative and local norming methods have in equitable identification of gifted students.

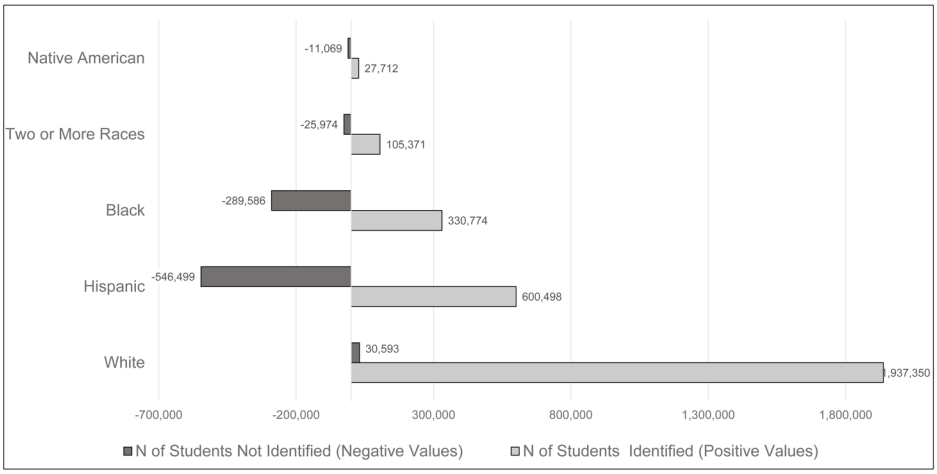
BACKGROUND

In this chapter the term gifted is used to describe students who are able to achieve advanced academic skills given the opportunity to learn. This is consistent with definitions of gifted and talented students provided by the U.S. Federal government and the National Association for Gifted Children (NAGC). The Federal government's definition is "Students, children, or youth who give evidence of high achievement capability in areas such as intellectual, creative, artistic, or leadership capacity, or in specific academic fields, and who need services and activities not ordinarily provided by the school in order to fully develop those capabilities." [Title IX, Part A, Definition 22. (2002)]. NAGC documents describe gifted students as those whose ability is significantly above the norm when compared to others of the same age and provide three points that are especially important. First is a distinction between gifted (i.e., a very smart person) and talented (i.e., a student with considerable knowledge and skills). Second is the view that gifted and talented students can be found in all racial, ethnic, and cultural groups, and across socioeconomic levels. Third, gifted and talented students need access to appropriate educational opportunities to reach their potential. These three points are central to the purpose and aims of this chapter, especially as it relates to the racial and ethnic composition of programs for gifted students.

Under-Representation

Black, Hispanic and Native American students as well as those learning English have been, and remain, considerably underrepresented in gifted education (Ford, 2013; Ford, Grantham, & Whiting, 2008; Office for Civil Rights (OCR), 2004, 2006, 2009, 2011, 2012; Yaluma & Tyner, 2018). For example, recent findings suggest that Black and Hispanic students were underrepresented in 47 states and Native

Figure 1. Numbers of Students Identified for Gifted Education by Race and Ethnicity.



American students were underrepresented in 43 states (Yoon & Gentry, 2009). The 2018-2019 State of the State in Gifted Education (NAGC, 2020) report found that Black and Hispanic students were grossly under-represented in 17 states as were Native American students. The same report also concluded that ELL students encompassed 2% or less of the gifted population in 12 states (NAGC, 2020). In addition, socioeconomic status (e.g., parental education) is associated with disproportionalities in gifted education (Callahan, 2005; List & Dykeman, 2019). The 2014-2015 State of the States in Gifted Education reported that low-income students make up 51% of the population but comprise as low as 1% in gifted education in some states (NAGC, 2015). How many students do these percentages represent?

The *National Center for Educational Statistics* provides statistics on the numbers of students in public schools throughout the US and the numbers of students identified as gifted and talented from 1999 to present and estimated to 2027. This allows for the calculation of the differences between the number of students *identified* as gifted and the number of *potentially* gifted students. Although NAGC suggests that gifted students represent the top 10% of the student population, we used a more conservative 8% (92nd percentile) value to compute the number of students who could be identified.

Our results summarized in Figure 1 shows that approximately 330,000 Black, 600,000 Hispanic, 40,000 Native Americans and 130,000 students with two or more races could have been identified as gifted and were not. Importantly, whites were identified at a rate slightly higher than the US population percentage. Hispanics and Blacks were under identified by about 50%, Native Americans 70%, and those having two or more races 80%. The total number of non-white students in K-12th grade who could reasonably have been, but were not, granted admission to gifted educational programs in the US today is approximately 873,000. This is an ongoing historical problem that must and can be addressed.

The large number of students of color who could have been identified for gifted programs but were not illustrates the failure of the methods used in the US. This is an important issue for educators and especially a critical social justice issue that has profound implications for society as a whole. There is no advantage to society to exclude so many smart students from the education they deserve particularly given that non-white student represent slightly more than the white population of public-school students

(Domenech, Sherman, & Brown, 2016). This historical problem is too severe to be ignored especially given that solutions are available if educators can overcome the historical impact of the tests, procedures and expectations that have dominated the field of gifted education for decades.

ADDRESSING UNDER-REPRESENTATION OF STUDENTS OF COLOR

Have you ever wondered what influence the content of ability tests has on the identification of gifted and talented students? Why do the tests contain verbal, nonverbal and quantitative test questions? Were the tests built on a specific theory of intelligence? Who built the first editions, when were they originally constructed and why? Do these tests measure ability in a fair and equitable way? These are, of course, critical questions because ability tests play a central role in the identification of gifted and talented students.

IQ Tests and Identification of Gifted Students

Individually administered intelligence tests such as the Stanford-Binet and the Wechsler scales have been used to identify gifted students for decades. Today, according to the results of a 2019 survey (Kurtz, Harwin, Chen & Furuya) the most widely used tests include, in descending order, the CogAT, Wechsler, Naglieri NAT, Woodcock, ITBS, Otis-Lennon, and Stanford-Binet. All these tests, except the Stanford-Binet were largely influenced by the efforts of the US Army Mental Testing program which produced the Army Alpha (verbal and quantitative) and Army Beta (nonverbal) tests (Yoakum & Yerkes, 1920). The past and current editions of all these ability tests contain subtests that fall into three traditional categories of verbal, nonverbal and quantitative.

It is important to consider the degree to which verbal and quantitative test questions are appropriate for diverse populations. One way to examine this is to determine if these tests yield different scores for diverse populations and if so, what is the impact of those scores on the composition of students who are accepted into programs for gifted students. To address this issue, we will present current research on race and ethnic differences on ability tests that demand knowledge and contrast those findings with evidence from tests that demand very little knowledge. We start with an important distinction between test bias and test equity.

The Standards for *Educational and Psychological Testing* make a distinction between test bias and test equity. Test bias is examined by studying group differences in subtest and scale reliabilities, item characteristic curves, factorial invariance, and other statistical ways to study what is called psychometric bias. These analyses look at the way the test items, subtests and scales function. Test equity is different, and in our view, more important from the practitioner perspective. The Standards clearly state that even if, for example an IQ test, does not show evidence of psychometric bias that test could still be socially unjust. The important part of the discussion of equity rests on the educational background of the student. That is, if a person has had limited opportunities to learn the content used in an intelligence test, that test may be considered unfair because it penalizes students for not knowing the answers even if there is no evidence of psychometric test bias. This means that all students must have had an equal opportunity to learn the content which influences the scores on an intelligence test. The extent to which these psychometric issues influence scores has been studied.

Race and ethnic differences for tests commonly used in the educational and psychological setting are provided in Table 1. We organized these tests by the size of the differences and what emerged was a clear

distinction between tests that require knowledge and those that require little knowledge. For example, tests that require a student to define words, solve word analogies, have general information, solve math word problems, understand verbal directions, and orally explain their answers were associated with larger differences by race and ethnicity. Tests, such as the NNAT which is completely nonverbal and the two versions of the Cognitive Assessment System 2nd Edition (Naglieri, Das, & Goldstein, 2014a & b) as well as the Kaufman Assessment Battery for Children – 2nd Edition (Kaufman & Kaufman, 2004) yielded the smallest differences. The findings provide a clear distinction between tests that demand knowledge versus those that do not except for the study of the CogAT.

Whereas nonverbal NNAT test scores showed small differences across race and ethnicity, Carman, Walther and Bartsch (2018) found that Black and Hispanic students earned scores on the CogAT7 nonverbal scale that were 9 Standard Age Score points less than whites. We suggest that the difference between these two nonverbal tests is likely the result of lengthy verbal instruction used for the CogAT7 in contrast to the NNAT which uses nonverbal pictorial directions. In total, these results clearly suggest that regardless of the test bias results, the measures that require knowledge are much less equitable. Equity has also been studied by measuring identification rates by race and ethnicity.

Table 1. Mean Score Differences in Intelligence Test Scores by Race & Ethnicity.

	Race	Ethnicity
Tests that require knowledge		
Otis-Lennon School Ability Test (school system)	13.6	
Stanford-Binet IV (normative sample)	12.6	
WISC-V (normative sample)	11.6	9.1
WJ- III (normative sample)	10.9	10.7
CogAT7 (Nonverbal scale)	11.8	7.6
WISC-V (statistical controls normative sample)	8.7	5.4
Tests that require minimal knowledge		
KABC-2 (matched samples)	5.0	
CAS-2 (normative sample)	6.3	4.5
CAS-2 (statistical controls normative sample)	4.5	1.8
NNAT (matched samples)	4.2	2.8
CAS2: Brief (normative samples)	2.0	2.8

Citations: For the Otis-Lennon School Ability Test by Avant and O'Neal (1986); Stanford-Binet IV from Wasserman & Becker (2000); Woodcock-Johnson III race differences from Edwards & Oakland (2006) and ethnic differences from Sotelo-Dynega, Ortiz, Flanagan & Chaplin (2013); CogAT7 from Carman, Walther and Bartsch (2018); WISC-V from Kaufman, Raiford & Coalson (2016); Kaufman Assessment Battery for Children-II from (Lichenberger, Sotelo-Dynega & Kaufman, 2009); CAS-2 and CAS2:Brief from Naglieri, Das & Goldstein, 2014a & 2014b; Naglieri Nonverbal Ability Test (Naglieri & Ronning, 2000).

Average score differences by race and ethnicity are helpful, but when considering the impact tests have on gifted identification rates it is important to ask the question: How many students are identified using a particular test? This question was studied using the Naglieri Nonverbal Ability Test (NNAT; Naglieri, 1997) by Naglieri and Ford (2003) for Black, Hispanic and White students. Their sample of

20,270 students (grades K–12) was similar to the U.S. population on the demographic variables used during norming of the test. The results were clear; similar percentages of Black (5.1%), Hispanic (4.4%) and White (5.6%), students had a standard score of 125 (95th percentile rank). These findings showed that the NNAT was an equitable way to find students for gifted education services. A second study of identification rates was conducted by Melissa Dayle Durtschi (2019). She examined the relationships between race and identification rates for nearly 3,000 students using the 90th and 95th percentile scores from the Cognitive Abilities Test 7 (CogAT7). She concluded that the significant difference between racial identification rates using CogAT7 scores resulted in under-representation of Black and Hispanic students.

Naglieri and Ford (2005) provide another example of how test items can pose a problem for those with limited reading skills. They evaluated the reading levels required for the Sentence Completion test on the CogAT Form 6 which is taken by children in grades 5 and 6. Using the Flesch-Kincaid Grade Level method (Flesch, 1948) they determined that the readability of the Sentence Completion test items was a grade level score of 6.1 with readabilities of the individual items ranging from grade 3.7 to 10.4. Children with poor reading skills, therefore, are likely to earn low scores on this kind of a test because of the reading level demands of this measure of ability.

In summary, these research studies show a clear pattern related to the amount of knowledge required by the various tests; those tests that require knowledge yield large differences across race and ethnicity which impacts the numbers of students of color who qualify for gifted educational services. Given the implications of these important findings the question becomes what characteristics should an equitable test of ability have? The answers to this question follow.

MEASURE THINKING NOT KNOWING TO ACHIEVE EQUITY

The ability tests used as part of the process of identifying gifted and talented students should measure a student's ability to think through problems to arrive at the answer. That is, ability tests inform us about how well a student can think and learn (i.e., a gifted student). In contrast, achievement tests tell us how much information a student has acquired from the educational curriculum (i.e., a talented student). This distinction is critical to equitable assessment, yet many ability test *instructions* demand listening comprehension, questions demand *knowledge* and in the case of individually administered ability tests, a verbal *response*.

Test Instructions

The administration instructions of any test must be written in a manner that ensures a student clearly understands what to do so that the student who fails a test question did so because they could not solve the problem not because they did not understand the instructions. For example, if the directions require knowledge of words that students may not have the opportunity to learn, then the ability test score will be contaminated by knowledge. Sometimes verbal comprehension demands of test directions is influenced by the use of verbal concepts. For example, a simple instruction such as "Find the big circle under the number 10" demands that the student knows concepts of size (e.g., big vs little), shape (e.g., circle vs square), orientation (e.g., under vs over) and numbers. Cummings and Nelson (1980) found that understanding basic concepts was necessary in group administered achievement test directions and that

students might not have mastered those concepts at the ages for which the tests were intended. Similarly, Gill, Moorer-Cook, Armstrong and Gill (2012) found that the grammatical and semantic aspects of test directions demand listening comprehension skills that students may not have. Randall, Engle, Carullo, Collins (2015) found that working memory capacity impacted students' ability to recall directions presented orally. All these factors can influence the score a student receives on any test, therefore, an ideal test would not require verbal instructions.

Test Questions

It is common among most ability tests used to identify gifted students to include questions that demand knowledge in general and *in particular,* knowledge of English. This includes test questions that require knowledge of vocabulary, word analogies, general information, and math word problems. Tests such as the Wechsler, CogAT and OLSAT all contain items like these which confounds the measurement of ability with what the student knows. Naglieri (2008) described the similar content in group tests of ability and achievement and amply demonstrated although these tests are intended to measure two different things many ability tests are confounded by knowledge. This is an important factor that is related to race, ethnic and socioeconomic differences (Naglieri & Otero, 2017) as shown above.

The Standards for Educational and Psychological Tests clearly state that the "extent to which individuals have had exposure to …knowledge that affords them the opportunity to learn the content and skills targeted by the test has several implications for the fair and valid interpretation of test scores (p. 56)". The Standards continue saying that "inequities in school resources available to students from traditionally disadvantaged groups, for example, racial, ethnic, language, and cultural minorities and rural students… affect the quality of education received … [and] the validity of the inferences about student ability … may be compromised (p. 56)". This is not a new problem, in fact roughly 100 years ago Pintner (1923) wrote "A good intelligence test must avoid as much as possible anything that is commonly learned (p. 61). In our view, an ideal test would be as free from knowledge acquired in school as possible.

Verbal Expression

Any time a student is required to provide a verbal response to a question, this can introduce a potential problem for equitable assessment. Individually administered ability tests that have verbal and quantitative test questions typically require students to explain, for example, in what way a fly and a tree are the same or provide a definition of words. This approach to assessment puts at disadvantage those with limited verbal knowledge and skills and provides an advantage to students who have had ample educational opportunities. The score a student receives on an ability test should not be influenced by the student's expressive language skills.

One Solution

Wechsler and Naglieri (2006) first introduced the use of test instructions that use pictures rather than verbal statements from an administrator in the Wechsler Nonverbal Scale of Ability. Two years later this method was incorporated into the second edition of the Naglieri Nonverbal Ability Test (Naglieri, 2008). These two tests also had minimal knowledge requirements and no verbal responses were required of the student. The success of this approach as it relates to equitable identification of gifted students is

perhaps best illustrated by the results of a study conducted by Card and Giuliano (2015). They found that using the NNAT2 as a universal assessment tool was associated with a marked increase in the identification of black and Hispanic students, as well as those from low socioeconomic background, for gifted programs. This finding is consistent with research on the various editions of the NNAT (e.g., Naglieri & Ronning, 2000; Naglieri & Ford, 2003) which showed that measuring thinking apart from knowing led to equitable assessment. The next question, however, is "can verbal and quantitative tests also be constructed in such a way that they measure general ability in an equitable manner?" Our answer is yes, it can be accomplished, and our solutions will be described next.

The Naglieri Verbal, Nonverbal & Quantitative General Ability Tests

The Naglieri General Ability Tests™ (Naglieri, Brulles & Lansdowne, 2021) were explicitly designed to meet the demand for equitable assessment using verbal, nonverbal and quantitative test content and eliminating the need for verbal instructions and verbal responses and minimizing the role of knowledge. Each of the three tests measure general ability using verbal (Naglieri-V™; Naglieri & Brulles, 2021), nonverbal (Naglieri-NV™; Naglieri, 2021), and quantitative (Naglieri-Q™; Naglieri & Lansdowne, 2021) test questions. These three separate tests are typically administered to groups but can be given individually using the online or paper formats. They all have questions presented using diagrams, numbers and pictures. Importantly, the individual items were carefully written so that minimal academic knowledge is needed, the questions can be solved using any language, cultural influences were minimized and no oral response is required.

Test Content

The three tests were specifically designed to measure thinking (i.e., general ability) with minimum influence of knowing. For example, the verbal test questions require a student to recognize a verbal concept represented by five pictures to determine how the sixth picture does not represent that same concept. The quantitative test questions require a student to detect the pattern(s) among numbers, symbols, and numerical sequences all of which only requires very basic math. The nonverbal test consists of questions which require a student to notice how shapes can form patterns when arranged in sequences, spatial orientations, and other distinguishing characteristics to arrive at the correct answer. All three of these tests and procedures used to administer them reflect how well a student can solve the test questions with minimal impact of extraneous variables.

Test Instructions and Responses

Instructions to the students are presented using animated videos when taken online or a four-frame pictorial version of the video when the tests are administered on paper. This method eliminates the need for comprehension of verbal instructions because the students are *shown* what they need to do rather than having someone *verbally tell* them what they are to do. This approach eliminates demands on memory, especially because students can play the instructions as many times as needed. In order to eliminate the role of verbal expression in the tests, students respond using a multiple-choice format. These methods of test construction were used so that general ability could be measured with tests that vary in content

Figure 2. Example of Naglieri Verbal Test Item

and can be solved using any language. Each of the three tests will be more fully explained in the sections that follow.

Verbal

The verbal version of the Naglieri General Ability Tests (Naglieri-V; Naglieri & Brulles, 2021) measures general ability using six pictures of ordinary objects five of which represent a verbal concept. This test was modeled after an approach described by A. R. Luria (1982) to evaluate verbal conceptual thinking. Luria stated that language involves, "a complex system of codes (p. 29)" where, "every word designates a thing, an attribute, an action or a relationship (p 34)." The task, referred to as superfluous fourth, demands that a subject reason and identify which word does not belong with the others, for example, rose, daisy, stem, tulip. Rather than use written or spoken words, items were constructed using pictures of easily recognizable objects typically seen and common across cultures. The task is simply to choose which of the six pictures does not represent a verbal concept that is shared by the other five. The test questions demand close examination of the relationships among all the pictures especially the attributes of the pictures (color, use, function, orientation, etc.), consideration of possible connections among the objects pictured and verbal reasoning to identify which five of the six go together. The test questions can be solved regardless of the language(s) the student knows.

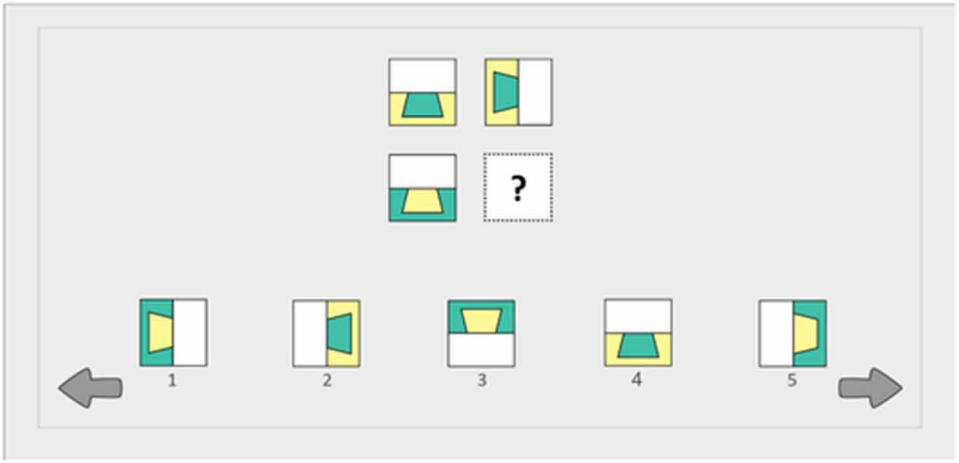
Figure 2 illustrates how a typical item on the verbal test was designed so that it could be solved using any language. The students' task is to detect the similarities among five of the six images. For example, some of the objects are round, have a handle, are similar in color, but only one is filled; so, the concept is empty. By varying the colors, shapes, and selection of pictures a nonverbal presentation of objects that represent a verbal concept can be detected using any language. The items become increasingly difficult as the complexity of the items increases.

Nonverbal

The nonverbal version of the Naglieri General Ability Tests (Naglieri-NV; Naglieri, 2021) measures general ability using geometric shapes that vary in color and orientation. The shapes may be arranged in sequences, vary in orientations, or change across a 2X2 or 3X3 matrix. All these attributes must be fully understood to determine the correct answer. The student's task is to determine the logic underlying the relationships among the parts and to identify which of five options is the answer. As the number of variables increases so too does the difficulty of the items.

Figure 3 provides an example of a typical nonverbal test question. The shapes change across the rows in color and the shapes' colors change in orientation across the columns. The student's complete understanding of the logic behind the changes in the shapes is needed to identify the correct answer (1). These items become increasingly difficult as the number of variables increases (e.g., color, location, shape, orientation, etc.), the configuration increases from a 2 by 2 to a 3 by 3, and the difficulty of the options increases.

Figure 3. Example of Naglieri Nonverbal Test Item



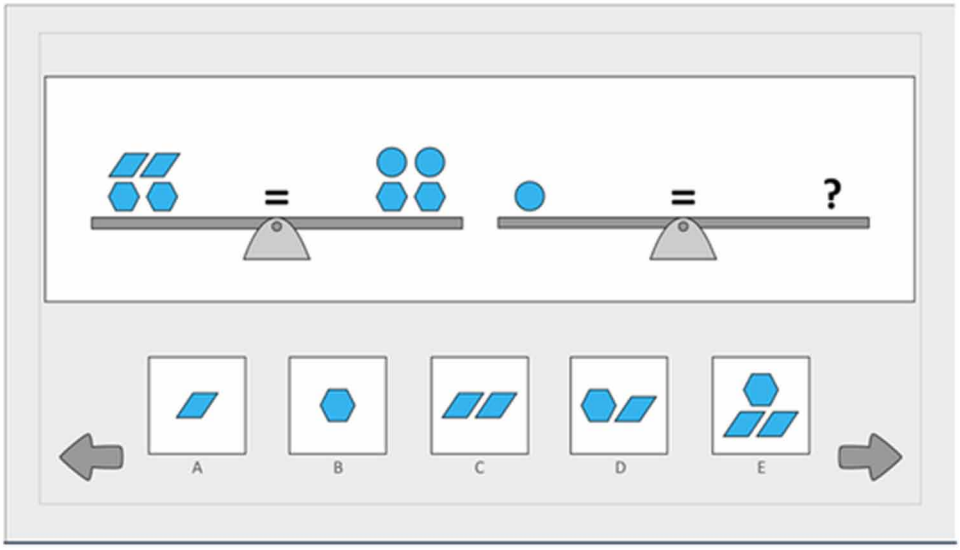
Quantitative

The quantitative version of the Naglieri General Ability Tests (Naglieri-Q; Naglieri & Lansdowne, 2021) measures general ability using numbers and symbols. Like the verbal and nonverbal test questions, the student must understand the relationships among the information provided, in this case numbers and symbols, to identify the correct answer. The questions involve equivalency of simple quantities, numerical sequences, numbers arranged in a matrix where minimal mathematical knowledge and calculation is required.

Figure 3 illustrates a typical quantitative test item that requires the student to notice two parallelograms and two circles balance each other as do the two hexagons on each side. Logically, one circle would be equal to one parallelogram. These items test a student's ability to understand relationships and patterns among shapes and numbers as is the case with the verbal and nonverbal test items. The test questions

become more difficult as the number of variables and reasoning becomes more complex. Importantly, math word problems are not used.

Figure 4. Example of Naglieri Quantitative Test Item



General Ability Using Verbal, Nonverbal and Quantitative Tests

The three tests of general ability were explicitly designed to vary by item content but at the same time measure general ability which we describe as the student's ability to fully understand relationships among pictures, shapes, or numbers to select an answer. The test questions, much like many academic tasks, demand that the students comprehend the logical relationships among verbal, nonverbal spatial and mathematical information represented by pictures, shapes, and numbers. By eliminating the need for verbal comprehension in the test directions, creating items that can be solved using any language, and removing the need for an verbal response we suggest that this approach meets the Standards for Educational and Psychological Tests definition of equitable assessment. It is this topic that we address next.

Evidence for the Naglieri General Ability Tests

After creating tests that we hoped would improve equity, the next important step was to study the extent to which this goal was achieved. Research support for any new idea is always necessary, and as it relates to the Naglieri General Ability Tests, such data is particularly important because of the impact test scores can have on the lives of many students. Three large-scale studies of these tests were recently conducted. Selvamenan, et al., (2021) studied the pre-publication versions of Naglieri General Ability Tests: Verbal, Nonverbal, and Quantitative. Three separate samples of students in Pre-K through Grade 12 were obtained (N= 8,105). Each of these tests were administered on a computer and had all the characteristics previously discussed. The authors report trivial differences on the Verbal, Nonverbal, and Quantitative tests by race, ethnicity, gender, and parental education. These findings support our expectation that by

reducing the language and knowledge demands in the items as well as the directions and student responses equity across race, ethnicity, gender, and parental educational levels can be achieved.

The Identification Process

We have shown that the selection of a measure of ability can have an impact on the extent to which a test is equitable across gender, race, ethnicity and parental educational levels. In reality, the test results typically fit within a larger assessment process and therefore we will focus on some practical issues.

Screening and Identification

Identification of students for gifted educational programs has traditionally been accomplished in a variety of ways oftentimes as described by state regulations, historical precedent, group and/or individual assessment, screening then verifying and so forth. It is important to recognize that how students are evaluated is just as important as what tests or rating scale are used, how these measures are scored and combined, what kinds of scores are used (e.g., local or national norms, percentile ranks, etc.), is the goal to match the student to the curriculum or the curriculum to the student, and what effect do the procedures used to combine all the data (e.g., a matrix) have on the outcome. Peters, et al. (2014) note, "Tests and other assessment tools do not make decisions. They are just tools that provide stakeholders with important information on which to base their decisions" (p. 147). We will examine some of these issues related to equitable assessment in the sections that follow.

Lee, Ottwein & Peters (2020) suggest that gifted programs should meet the needs of those gifted learners for whom the regular curriculum and instruction may be inappropriate. When the curriculum does not satisfy the student's needs underperformance in the regular classroom, behavioral issues, or simply lack of motivation to perform can result. Students who are not performing well academically in school typically would not be recommended for possible placement in a gifted educational program. For this reason, universal assessment should be used to identify gifted (i.e., very smart) students.

Some school district administrators use various types of screening tools as a way to begin the identification process. However, screening tools, such as rating scales or very short tests, *decide* who should be given additional tests and will not provide equal opportunity for *all* students to demonstrate their ability. Administration of ability tests such as the ones described here can be given to all students, provide information that can be used for identification and eliminate the need for a screening test. When tests are administered to *all* students everyone has the opportunity to be identified. Another possible obstacle to equitable identification is how multiple sources of information are combined.

The logic used in the eligibility process for students to be accepted in a gifted education program is just as important as the information used in the assessment process. A good example is found in the McFadden v. Board of Education for Illinois School Dist. U-46 (2013; 984 F.Supp.2d 882). The district was composed of approximately 40% Hispanic students yet only 2% of students in the gifted program were Hispanic. Parents argued that the district failed to eliminate language barriers in violation of the Equal Education Opportunity Act and limited educational benefits for Hispanics, in violation of federal and state laws. The district required that all students accepted to the gifted program attain high scores on measures of ability that demanded knowledge of English and academic math skills (i.e., the CogAT6; Lohman & Hagen, 2001). The judge ruled that the tests that demanded knowledge of English excluded Hispanic students. He stated that "gifted children for whom English is a second language would likely

score lower on a [verbal] test than the nonverbal, culturally neutral Naglieri Nonverbal Ability Test"(p. 24). This case illustrates that having an equitable test as part of the identification process does not necessarily result in equitable identification unless the most equitable scores are valued higher than the academically laden measures that served as barriers.

National and Local Norms

Equitable representation is influenced by the content of the tests used in the identification process, how the tests are used (Peters, et al, 2016) and the definition of gifted and talented students. According to NAGC (2019) "Students with gifts and talents perform—or have the capability to perform—at higher levels compared to others of the same age, experience, and environment in one or more domains." The statement "compared to others of the same age" is very important because the comparison can be accomplished by a national or local point of reference.

Ability tests have been calibrated using national norms since the early 1900s. This means that the raw scores (i.e., the number of correct answers) a student earns is compared to a sample of same aged or grade students who were tested using the same test. The composition of that sample of same aged students (selected to represent the country) is what defines a national norm. The sample of students used to create a national norm must be large (i.e. thousands of students) and closely represent the country on the basis of critical variables such as age, sex, race, ethnicity, parental educational levels (socioeconomic status), and urban/rural settings (AERA, APA, NCME, 2014). Publishers and authors have a responsibility to fully describe the normative sample and especially its similarity to the country on those demographic variables. A test that is nationally normed, therefore, is based on the comparison of an individual to a group that represents the entire country.

A test that is locally normed provides a comparison of an individual to students in the same grade, school or school district. In this case the comparison group closely represents the local community on the basis of critical variables such as age, sex, race, ethnicity, parental educational levels (socioeconomic status), and urban/rural settings. The sample used to create a local norm should be reasonably large so that it adequately represents the local demographics. "Local norms are often useful in conjunction with published norms, especially if the local population differs markedly from the population on which published national norms were based. In some cases, local norms may be used exclusively" (AERA, APA, NCME, 2014; p. 196).

National and local norming procedures provide two different, but, equally important comparisons. First, the score a student earns is compared to a representative sample of students the same age across the country. Second, the student's score is compared to the local community. Both methods of calibrating scores are used in the field of gifted and talented education. The decision to use one or the other can be influenced by the demographic composition of the school and/or district. For example, when the school and/or district's population is economically, linguistically, ethnically, and/or racially diverse and differs markedly from the national population, local norms provide a comparison that more precisely reflects the characteristics of the community in which the students reside.

Example of Local Norm Calculations

To illustrate exactly how a local norming process can help identify the most capable students we created a fictitious sample of 94 cases and used the data to show how local norming works. This example is

based on the combination of four grade 2 classes and two different groups of grade 2 classes that could be from different parts of a community which vary on a number of demographic variables. Using local norms for each part of the community allows scores to be obtained that more precisely represent the specific community within which students reside. In this example, three separate local norms were created.

Local norms provided in Figure 5 were computed on the number of questions answered correctly (called Student Raw Score) on a test. Each student's Raw Score was used to rank all the students so that the case with the highest score obtained a rank of 1. The ranks are used to obtain the Local Percentile Ranks which are interpreted as follows: a student with a percentile score of 90 did as well as or better than 90% of the students in the reference group.

The first local norm was created using data from all four second grade classes in the district (N = 94). This comparison represents the diverse composition of the second-grade classes in the entire school district. The table includes values that identify which classroom each case was assigned to, a student identification number, the ranking for each student and their percentile score based on all 94 students in grade 2 in the district. Examination of these results show that case number 34 earned a raw score of 52, which is a local rank of 3 and a percentile score of 96. This means that raw score was equal to or better than 96% of the students in this group.

The second and third set of local norms was based on (a) two grade 2 classes from one part of the community (n = 46); and (b) two grade 2 classes (n = 48) from another part of the community. The local norms based on each of the pairs of second grade classes represent the two communities more precisely. These two groups include all those cases used in the total sample of 94. Note that the cases in all three of these groups were ranked according to the Student Raw Scores.

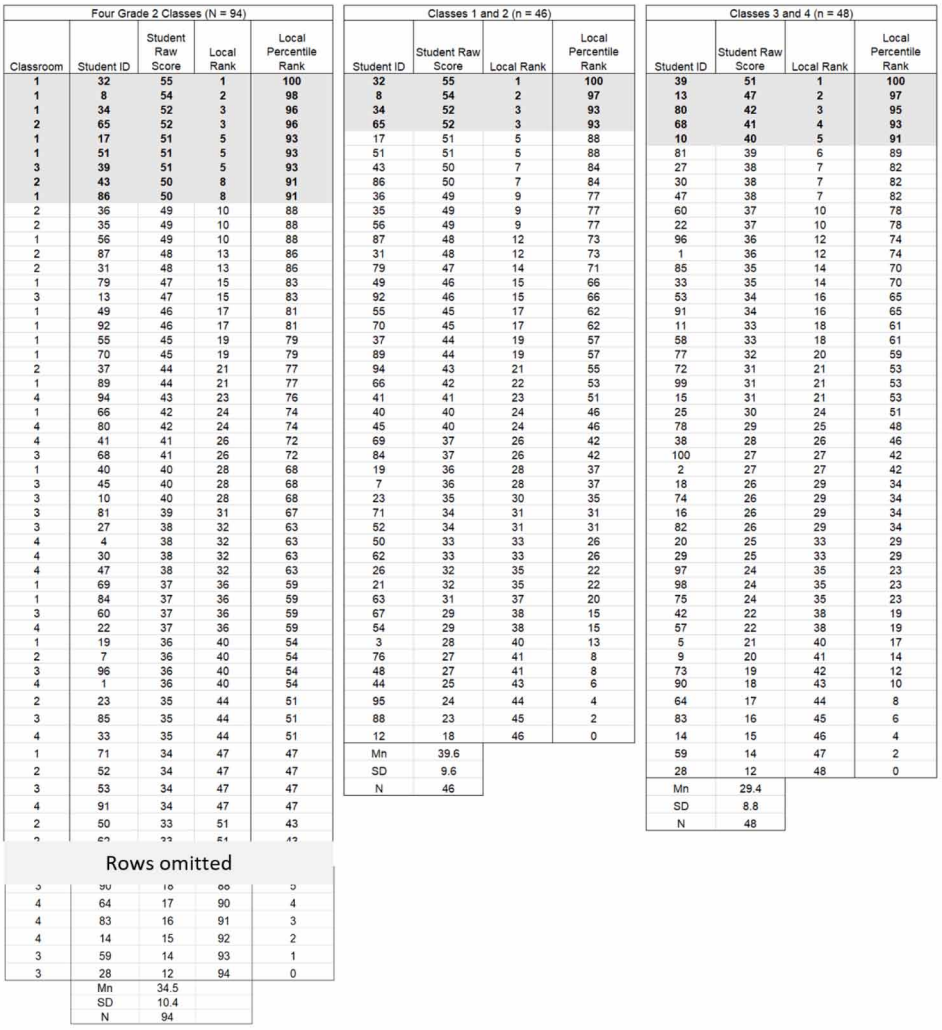
To understand the results, we need to begin by examining the average score usually referred to as the mean (Mn) in each group. Notice that the district mean is 34.5 correct out of 55 questions. When all the students in the four second grade classes are combined the group of students who scored above the 90th percentile are identified. Students with identification numbers 32, 8, 34, 65, 17, 51, 39, 43, and 86 all had a Local Percentile Rank of 90 or higher based on the comparison to all second graders in the school district.

The mean score for Classes 1 and 2 of 39.6 is higher than the district mean of 34.5. The group with percentile scores of 90 and above is composed of the top four students in the district local norm (32, 8, 34, & 65). Note, however, that their Local Percentile Rank scores are slightly different from what they were when compared to students in all four second grade classes.

The mean score for Classes 3 and 4 was 29.4, lower than the district mean. The students who earned scores at the 90th percentile or higher in this group consisted of Student ID numbers of 39, 13, 80, 68, and 10. One of these cases (number 39) also earned a score at the 90th percentile or higher using the district local norm. The other four cases who earned scores at or above the 90th percentile in relation to the group composed of classes 3 and 4 did not receive such scores using the district level local norms.

In summary, the local norming process is a way of calculating scores in relation to different reference points. Changing the reference group can change the Local Percentile Rank. These data illustrate how a student's score is dependent upon how many questions they get correct on a test and how well their score compares to different groups of others who took the same test. This is, however, only one example and outcome of using district or school based local norms, which will vary, by definition, by the composition of the groups.

Figure 5. Local Norms Example



Four Grade 2 Classes (N = 94)

Classroom	Student ID	Student Raw Score	Local Rank	Local Percentile Rank
1	**32**	**55**	**1**	**100**
1	**8**	**54**	**2**	**98**
1	**34**	**52**	**3**	**96**
2	**65**	**52**	**3**	**96**
1	**17**	**51**	**5**	**93**
1	**51**	**51**	**5**	**93**
3	**39**	**51**	**5**	**93**
2	**43**	**50**	**8**	**91**
1	**86**	**50**	**8**	**91**
2	36	49	10	88
2	35	49	10	88
1	56	49	10	88
2	87	48	13	86
2	31	48	13	86
1	79	47	15	83
3	13	47	15	83
1	49	46	17	81
1	92	46	17	81
1	55	45	19	79
1	70	45	19	79
2	37	44	21	77
1	89	44	21	77
4	94	43	23	76
1	66	42	24	74
4	80	42	24	74
4	41	41	26	72
3	68	41	26	72
1	40	40	28	68
3	45	40	28	68
3	10	40	28	68
3	81	39	31	67
3	27	38	32	63
4	4	38	32	63
4	30	38	32	63
4	47	38	32	63
1	69	37	36	59
1	84	37	36	59
3	60	37	36	59
4	22	37	36	59
1	19	36	40	54
2	7	36	40	54
3	96	36	40	54
4	1	36	40	54
2	23	35	44	51
3	85	35	44	51
4	33	35	44	51
1	71	34	47	47
2	52	34	47	47
3	53	34	47	47
4	91	34	47	47
2	50	33	51	43
Rows omitted				
4	64	17	90	4
4	83	16	91	3
4	14	15	92	2
3	59	14	93	1
3	28	12	94	0
	Mn	34.5		
	SD	10.4		
	N	94		

Classes 1 and 2 (n = 46)

Student ID	Student Raw Score	Local Rank	Local Percentile Rank
32	**55**	**1**	**100**
8	**54**	**2**	**97**
34	**52**	**3**	**93**
65	**52**	**3**	**93**
17	51	5	88
51	51	5	88
43	50	7	84
86	50	7	84
36	49	9	77
35	49	9	77
56	49	9	77
87	48	12	73
31	48	12	73
79	47	14	71
49	46	15	66
92	46	15	66
55	45	17	62
70	45	17	62
37	44	19	57
89	44	19	57
94	43	21	55
66	42	22	53
41	41	23	51
40	40	24	46
45	40	24	46
69	37	26	42
84	37	26	42
19	36	28	37
7	36	28	37
23	35	30	35
71	34	31	31
52	34	31	31
50	33	33	26
62	33	33	26
26	32	35	22
21	32	35	22
63	31	37	20
67	29	38	15
54	29	38	15
3	28	40	13
76	27	41	8
48	27	41	8
44	25	43	6
95	24	44	4
88	23	45	2
12	18	46	0
Mn	39.6		
SD	9.6		
N	46		

Classes 3 and 4 (n = 48)

Student ID	Student Raw Score	Local Rank	Local Percentile Rank
39	**51**	**1**	**100**
13	**47**	**2**	**97**
80	**42**	**3**	**95**
68	**41**	**4**	**93**
10	**40**	**5**	**91**
81	39	6	89
27	38	7	82
30	38	7	82
47	38	7	82
60	37	10	78
22	37	10	78
96	36	12	74
1	36	12	74
85	35	14	70
33	35	14	70
53	34	16	65
91	34	16	65
11	33	18	61
58	33	18	61
77	32	20	59
72	31	21	53
99	31	21	53
15	31	21	53
25	30	24	51
78	29	25	48
38	28	26	46
100	27	27	42
2	27	27	42
18	26	29	34
74	26	29	34
16	26	29	34
82	26	29	34
20	25	33	29
29	25	33	29
97	24	35	23
98	24	35	23
75	24	35	23
42	22	38	19
57	22	38	19
5	21	40	17
9	20	41	14
73	19	42	12
90	18	43	10
64	17	44	8
83	16	45	6
14	15	46	4
59	14	47	2
28	12	48	0
Mn	29.4		
SD	8.8		
N	48		

CONCLUDING COMMENTS

Advocates for underrepresented populations in gifted programs have long cried for change but change has been slow to happen. This chapter is about change. It provides a critical account of the procedures used to identify gifted and talented students, particularly students of color, low socio-economic status and non-English speakers. This is not a new problem; historical underrepresentation of students of color in programs for gifted students has been well documented in research and literature for many decades. Over 70 years ago, in the late 1950's Sputnik took flight and, out of fear of being surpassed as a global superpower, the US government committed schools to gifted education. In an effort to produce scientists, doctors, mathematicians and researchers, American students were identified as showing promise for accomplishing great things. An enormous amount of time, money and resources were devoted to "finding the best and the brightest'' students, and it is not surprising that most of these students were white,

came from affluent families and were already destined for, if not greatness, goodness. Seventy years is a long time to make things right and yet, little progress has been made toward equitably identifying gifted children and there continues to be disproportionate numbers of identified students who are poor, disenfranchised, speak a language other than English, are Black, Brown or Native American.

As described in this chapter, the critical issue lay in the way in which students are being identified for gifted services. The assessment is critically important, as it ultimately could alter a student's educational pathway, which can potentially change a student's life. These authors describe gifted students as those with high general ability (regardless of their academic skills) and talented students as those who have high achievement (advanced academics). This distinction has considerable implications for the types of tests used in the identification process. That is, five out of the six most widely used tests to identify gifted and talented students have characteristics that result in under-representation of Black, Hispanic, Native American and ELL students.

Considering the number of missing students of color and diversity in our gifted programs, clearly, historically prevalent, identification procedures are unequitable. Continuing to use the same approaches toward identification and expecting different outcomes will not lead to changes in the students we identify for gifted services, and perhaps more importantly, will not honor the intelligence of students who are unable to perform on typical IQ tests due to lack of experience, education and advocacy. This chapter/book is about challenging the status quo. Giftedness should have no barriers associated with race, socioeconomic status, language spoken, color of skin but it does and has for a very long time. We can do better and we must.

This chapter challenges educators to think differently, take risks for kids that have a spark, idiosyncrasy, passion, out-of-the-box idea regardless of the color of their skin, the education of their parent(s), the language they speak, and the enrichment of their early home life. These students deserve an enlightened approach and outlook. Their future may depend on educators identifying their cognitive and personal strengths along with their potential to learn.

They are worth it.

REFERENCES

AERA, APA, & NCME. (2014). Standards for Educational and Psychological Testing. Washington, DC: AERA.

Avant, A. H., & O'Neal, M. R. (1986, November). *Investigation of the Otis-Lennon School Ability Test to predict WISC-R Full Scale IQ for referred children*. *Annual Meeting of the Mid-South Educational Research Association*, Memphis, TN, United States.

Borland, J. H. (2009). Myth 2: The gifted constitute 3% to 5% of the population. Moreover, giftedness equals high IQ, which is a stable measure of aptitude. *Gifted Child Quarterly*, *53*(4), 236–238. doi:10.1177/0016986209346825

Callahan, C. M. (2005). Identifying gifted students from underrepresented populations. *Theory into Practice*, *44*(2), 98–104. doi:10.120715430421tip4402_4

Card, D., & Giuliano, L. (2015). *Can Universal Screening Increase the Representation of Low Income and Minority Students in Gifted Education?* NBER Working Papers 21519, National Bureau of Economic Research, Inc.

Carman, C. A., Walther, C. A., & Bartsch, R. A. (2018). Using the Cognitive Abilities Test (CogAT) 7 Nonverbal Battery to Identify the Gifted/Talented: An Investigation of Demographic Effects and Norming Plans. *Gifted Child Quarterly*, *62*(2), 193–209. doi:10.1177/0016986217752097

Cummings, J. A., & Nelson, R. B. (1980). Basic concepts in the oral directions of group achievement tests. *The Journal of Educational Research*, *73*, 259–261.

Domenech, D., Sherman, M., & Brown, J. L. (2016). *Personalizing 21st century education: A framework for student success*. Jossey-Bass.

Durtschi, M. D. (2019*). Inclusive Pathways to Gifted Education: Examining Gifted Referral Processes* [Unpublished doctoral dissertation]. University of Colorado, Denver, CO, United States.

Edwards, O. W., & Oakland, T. D. (2006). Factorial invariance of Woodcock-Johnson III scores for African Americans and Caucasians Americans. *Journal of Psychoeducational Assessment*, *24*(4), 358–366. doi:10.1177/0734282906289595

Engle, R. W., Carullo, J. J., & Collins, K. W. (1991). Individual Differences in Working Memory for Comprehension and Following Directions. *The Journal of Educational Research*, *84*(5), 253–262. doi: 10.1080/00220671.1991.10886025

Flesch, R. (1948). A new readability yardstick. *The Journal of Applied Psychology*, *32*(3), 221–233. doi:10.1037/h0057532 PMID:18867058

Ford, D. Y. (1998). The underrepresentation of minority students in gifted education: Problems and promises in recruitment and retention. *The Journal of Special Education*, *32*(1), 4–14. doi:10.1177/002246699803200102

Ford, D. Y. (2013). *Recruiting and retaining culturally different students in gifted education*. Prufrock Press.

Ford, D. Y., Grantham, T. C., & Whiting, G. W. (2008). Culturally and linguistically diverse students in gifted education: Recruitment and retention issues. *Exceptional Children*, *74*(3), 289–306. doi:10.1177/001440290807400302

Frasier, M. M., Garcia, J. H., & Passow, A. H. (1995). *A review of assessment issues in gifted education and their implications for identifying gifted minority students (RM95204)*. University of Connecticut, National Research Center on the Gifted and Talented.

Gill, C., Moorer-Cook, L., Armstrong, E. S., & Gill, K. (2012). The Ability to Follow Verbal Directions: Identifying Skill Levels and Measuring Progress. *Canadian Journal of Speech-Language Pathology and Audiology*, *36*, 234–247.

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman assessment battery for children* (2nd ed.). American Guidance Service.

Kaufman, A. S., Raiford, S. E., & Coalson, D. L. (2016). *Intelligent testing with the WISC-V*. John Wiley & Sons, Inc.

Kurtz, H., Lloyd, S., Harwin, A., Chen, V., & Furuya, Y. (2019). *Gifted Education: Results of a National Survey*. Education Week Research Center.

Lee, L. E., Ottwein, J. K., & Peters, S. J. (2020). Eight universal truths of identifying students for advanced academic interventions. In J. H. Robins, J. L. Jolly, F. A. Karnes, & S. M. Bean (Eds.), *Methods and materials for teaching the gifted* (5th ed., pp. 61–80). Prufrock Press.

Lichtenberger, E. O., Sotelo-Dynega, M., & Kaufman, A. S. (2009). The Kaufman Assessment Battery for Children—Second Edition. In J. A. Naglieri & S. Goldstein (Eds.), Practitioner's guide to assessing intelligence and achievement (pp. 61–93). John Wiley & Sons Inc.

List, A., & Dykeman, C. (2019). *The Relationship of Gifted Program Enrollment to Race/Ethnicity, Gender, and SES*. doi:10.31234/osf.io/zx396osf.io/zx396

Lohman, D. F., & Hagen, E. P. (2001). *Cognitive Abilities Test (Form 6)*. Riverside.

Luria, A. R. (1982). *Language and cognition*. Wiley.

McFadden v. Board of Education for Il, S. Dist. U-46. No. 05 C 760 United States District Court, N.D. Illinois, Eastern Division. Dec 1, 2009

Naglieri, J. A. (1997). *Naglieri Nonverbal Ability Test*. The Psychological Corporation.

Naglieri, J. A. (2008). *Naglieri Nonverbal Ability Test* (2nd ed.). Pearson.

Naglieri, J. A. (2008). Traditional IQ: 100 Years of Misconception and its Relationship to Minority Representation in Gifted Programs. In J. VanTassel-Baska (Ed.), Critical Issues in Equity and Excellence in Gifted Education Series Alternative Assessment of Gifted Learners (pp. 67-88). Waco, TX: Prufrock Press.

Naglieri, J. A. (2008). *Naglieri Nonverbal Ability Test* (2nd ed.). Pearson.

Naglieri, J. A. (2021). *Naglieri Ability Test: Nonverbal*. Multi-Health Systems.

Naglieri, J. A., & Brulles, D. (2021). *Naglieri Ability Test: Verbal*. Multi-Health Systems.

Naglieri, J. A., Brulles, D., & Lansdowne, K. (2021). *Naglieri General Ability Tests*. Multi-Health Systems.

Naglieri, J. A., Das, J. P., & Goldstein, S. (2014a). *Cognitive Assessment System* (2nd ed.). ProEd.

Naglieri, J. A., Das, J. P., & Goldstein, S. (2014b). Cognitive Assessment System Second Edition: Brief. Austin, TX: ProEd.

Naglieri, J. A., & Ford, D. Y. (2003). Addressing underrepresentation of gifted minority children using the Naglieri Nonverbal Ability Test (NNAT). *Gifted Child Quarterly*, *47*(2), 155–160. doi:10.1177/001698620304700206

Naglieri, J. A., & Ford, D. Y. (2005). Increasing minority children's participation in gifted classes using the NNAT: A response to Lohman. *Gifted Child Quarterly*, *49*(1), 29–36. doi:10.1177/001698620504900104

Naglieri, J. A., & Lansdowne, K. (2021). *Naglieri Ability Test: Quantitative*. Multi-Health Systems.

Naglieri, J. A., & Otero, T. M. (2017). *Essentials of CAS2 Assessment*. Wiley.

Naglieri, J. A., & Ronning, M. E. (2000). Comparison of White, African-American, Hispanic, and Asian Children on the Naglieri Nonverbal Ability Test. *Psychological Assessment*, *12*(3), 328–334. doi:10.1037/1040-3590.12.3.328 PMID:11021156

National Center for Education Statistics. (2013). *Children 3 to 21 years old served under Individuals with Disabilities Education Act (IDEA), Part B, by race/ethnicity and type of disability: 2010–11 and 2011–12*. Retrieved from https://nces.ed.gov/programs/digest/d13/tables/dt13_204.50.asp

Olszewski-Kubilius, P., & Corwith, S. (2018). Poverty, academic achievement, and giftedness: A literature review. *Gifted Child Quarterly*, *62*(1), 37–55. doi:10.1177/0016986217738015

Peters, S. J., & Engerrand, K. G. (2016). Equity and excellence: Proactive efforts in the identification of underrepresented students for gifted and talented services. *Gifted Child Quarterly*, *60*(3), 159–171. doi:10.1177/0016986216643165

Peters, S. J., Matthews, M., & McCoach, B. (2014). *Beyond gifted education: Designing and implementing advanced academic programs*. Prufrock Press Inc.

Pintner, R. (1923). *Intelligence testing*. Holt.

Selvamenan, M., Paolozza, A., Solomon, J., Naglieri, J. A., & Schmidt, M. T. (2020). *Race, Ethnic, Gender, and Parental Education Level Differences on Verbal, Nonverbal, and Quantitative Naglieri General Ability Tests: Achieving Equity*. Manuscript submitted for publication.

Sotelo-Dynega, M., Ortiz, S. O., Flanagan, D. P., & Chaplin, W. F. (2013). English language proficiency and test performance: An evaluation of bilingual students with the Woodcock-Johnson III Tests of Cognitive Abilities. *Psychology in the Schools*, *50*(8), 781–797. Advance online publication. doi:10.1002/pits.21706

Wasserman, J. D., & Becker, K. A. (2000, August). *Racial and ethnic group mean score differences on intelligence tests*. *Annual meeting of the American Psychological Association*, Washington, DC.

Wechsler, D., & Naglieri, J. A. (2006). *Wechsler Nonverbal Scale of Ability*. Pearson.

Yaluma, C. B., & Tyner, A. (2018). Are U.S. Schools Closing the "Gifted Gap"? Analyzing Elementary and Middle Schools' Gifted Participation and Representation Trends (2012–2016). *Journal of Advanced Academics*. Advance online publication. doi:10.1177/1932202X20937633

Yoakum, C. S., & Yerkes, R. M. (1920). *Army Mental Tests*. Henry Holt and Company. doi:10.1037/11054-000

Yoon, S. Y., & Gentry, M. (2009). Racial and ethnic representation in gifted programs: Current status of and implications for gifted Asian American students. *Gifted Child Quarterly*, *53*(2), 121–136. doi:10.1177/0016986208330564