Selecting for Ethnically Diverse Children Who May Be Gifted Using Raven’s Standard Progressive Matrices and Naglieri Nonverbal Abilities Test

Joan D. Lewis, Stephanie S. DeCamp-Fritson, Jean C. Ramage, Max A. McFarland, & Teara Archwamety

The identification of ethnically diverse students who are gifted has become a topic of great concern in the educational arena. Over the past few decades there has been increasing concern about the under-representation of ethnically diverse students in programs for gifted and talented students (e.g., Chambers, Barron, & Sprecher, 1980; Hadaway & Marek-Shroer, 1992; Karnes & Whorton, 1988; Mills & Tissot, 1995; Stephens, Kiger, Karnes, & Whorton, 1999; Shaunessy, Karnes, & Cobb, 2004).

Numerous programs have been developed to provide assistance to ethnically diverse students experiencing academic difficulties, yet few programs have focused on identifying and providing appropriately high level instruction for ethnically diverse children who are gifted. An immeasurable amount of talent is left unrecognized and under-developed as these children continue to be excluded from many gifted programs. According to Ford (1996), African American, Hispanic American, and Native American gifted students may be under-representation by as much as 50%.

Many explanations have been suggested for this under-representation of ethnically diverse students in gifted and talented programs. The screening and identification process for high-ability learners has come under scrutiny in the search for answers. In particular, it has been suggested that traditional measures of cognitive abilities are biased against certain groups of students (e.g., Chambers et al., 1980; Hadaway & Marek-Shroer, 1992; Johnsen, 2004; Karnes & Whorton, 1988; Mills & Tissot, 1995; Stephens et al., 1999).

The traditional measures of cognitive abilities in question include IQ tests, standardized achievement tests, and aptitude tests. Chambers et al. (1980) and Stephens et al. (1999) report that many of these tests are culturally loaded verbal assessment devices that do not take into consideration the colloquial language used by many culturally different children.

A second explanation for the under-representation of ethnically diverse children in gifted education programs is that many of these children may not have acquired the skills necessary to be successful in demanding programs for gifted students (Mills, Stork, & Krug, 1992; Mills & Tissot, 1995). Many of the under-represented students can be considered educationally disadvantaged as a result of educational, linguistic, cultural, and other environmental factors, causing disparity in test performance. These differences could be a result of inconsistencies in skill acquisition at the time of the test, and not test bias (Mills et al., 1992; Mills & Tissot, 1995).

Inadequate academic preparation may be the reason many ethnically diverse children who may be gifted fail to be identified with traditional forms of assessment. These children may not have acquired the knowledge base necessary to be identified for programs that build upon previously learned academic skills. According to the Javits federal definition found in the No Child Left Behind Act of 2001, giftedness refers to 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 or activities not ordinarily provided by the school in order to fully develop those capabilities. (Pub. L. No. 107-110, Title IX)

Since gifted learners can be found in any population, there are ethnically diverse students who could possess the cognitive skills necessary to succeed in programs for gifted learners; however, if they have any academic skill deficiencies, they may need interventions to help them develop their potential (Karnes & Whorton, 1988; Mills et al., 1992; Mills & Tissot, 1995).

Ethnically diverse students who are gifted can be successful in programs for academically talented students if they are first prepared for the program. First they must be identified. This could mean using non-traditional measures to locate these students and then “providing them with an intervention that would help them develop their potential by strengthening their academic skills and higher-level reasoning” (Mills & Tissot, 1995, p. 210).

This non-traditional form of selection would need to take into account some of the reasons for the under-representation of ethnically diverse students, and should cast a wider net by allowing a larger number of students to be provided with interventions. This type of approach has been shown to be successful with culturally diverse students (Mills et al., 1992).

These explanations for the under-representation of ethnically diverse students in gifted programs suggest that alternative methods of selection that are not based on acquired academic skills or verbal abilities may be necessary. For example, students could be assessed using universal reasoning and problem-solving skills. Ideally, this form of assessment would be free of bias against race, gender, ethnicity, and socioeconomic status. Many nonverbal tests have been constructed in an attempt to reduce such bias (e.g., Bracken & McCallum, 1997; Brown, Sherbenou, & Johnsen, 1997; Catell & Catell, 1965; Hammill, Pearson, & Wiederholdt, 1997; Kaufman & Kaufman, 1990; Naglieri, 1986; Raven, 1958; Raven, Raven, & Court, 1998).

Research investigating the effectiveness of nonverbal abilities tests has become increasingly popular with the growing recognition of the need for reduced-biased testing. Numerous studies have been conducted on the usefulness of these devices.
in selecting for students who are gifted (e.g., Karnes & McGinnis, 1994; Karnes & Whorton, 1988; Lewis, 1999; Mills, Ablard, & Brody, 1993; Mills & Tissot, 1995; Naglieri, & Ford, 2003; Shaunessy, Karnes, & Cobb, 2004; Stephens et al., 1999).

One form of nonverbal assessment that has been suggested by many researchers as an alternate or supplementary measure in identifying gifted students from culturally diverse backgrounds is the Raven’s Standard Progressive Matrices (Karnes & Whorton, 1988; Mills & Tissot, 1995; Richmond, 1987; Shaunessy et al., 2004; Stephens et al., 1999). The Raven’s is generally regarded as a nonverbal measure of fluid intelligence (Mills & Tissot, 1995).

The test developers claim the test measures higher-level thought processes, including the ability to reason by analogy and the ability to become more efficient by learning from immediate experience (Raven et al., 1998). With the Raven’s it is possible to learn from the easier items in order to improve performance on the more difficult items, yielding an index of intellectual efficiency that has many implications for identifying culturally diverse students who may be gifted (Mills & Tissot, 1995; Raven et al., 1998).

Mills and Tissot (1995) found the Raven’s identified a significantly greater percentage of ethnically diverse students who were gifted, many of whom were low-achieving students, than the School and College Ability Test, a more traditional measure of academic aptitude. Stephens et al. (1999) found that when compared with the Naglieri Nonverbal Abilities Test and the Culture-Fair Intelligence Test (CFIT, Cattell & Cattell, 1965); both nonverbal assessment devices, the Raven’s identified the largest number of ethnically diverse students scoring at the 80th percentile or higher.

Shaunessy et al. (2004) reported similar results while Lewis (1999) found that the Raven’s and CFIT revealed similar numbers of culturally different students although each test discovered some students the other did not. The results of these studies indicate that the Raven’s Standard Progressive Matrices may be an effective means of screening ethnically diverse gifted students.

A second nonverbal abilities test that may be useful in selecting for ethnically diverse students who may be gifted is the Naglieri Nonverbal Abilities Test (NNAT). The NNAT administration manual indicates that the test items were “developed to assess ability without requiring the student to read, write, or speak” (Naglieri, 1996, p.3). Test items were selected in a manner that attempted to have no bias against race, gender, or ethnicity. The NNAT allows students to use reasoning and problem solving skills, not verbal abilities.

Naglieri (1996) states in the directions for administering and scoring the Naglieri Nonverbal Abilities Test:

NNAT is appropriate for students from diverse cultural and language backgrounds, including students whose school performance may be poor because of limited proficiency in English and gifted and talented students who are either non-English speakers or are just learning English. NNAT is also designed for fair assessment of socially or economically disadvantaged students. (p.3)

The NNAT may be a useful tool in selecting for ethnically diverse students who may be gifted (Naglieri, 1996). Although the NNAT is one of the most recent nonverbal abilities tests, there have been a number of recent studies comparing its ability to screen for gifted students with the Raven’s and the CFIT. Although Lewis (1999), Shaunessy et al. (2004), and Stephens et al. (1999) all reported that the NNAT found the fewest potentially gifted students at or above the 80th percentile with different groups of culturally diverse students, Naglieri and Ford (2003) stated that they found similar proportions of students in each cultural group.

Research conducted on its predecessor, the Matrix Analogies Test-Short Form (MAT) suggests that the MAT correlates highly with the Wechsler Intelligence Scale for Children-Revised (r=.52) (Karnes & McGinnis, 1994). A study by Prewett (1995) indicates that the correlation between the MAT and the Wechsler Intelligence Scale for Children-Third Edition (WISC-III) is significant (r=.67).

These findings appear to contradict claims made by the NNAT, as the WISC-R and the WISC-III are traditional forms of assessment that have been criticized for their failure to identify ethnically diverse students as gifted (Chambers et al., 1980; Hadaway & Marek-Shroer, 1992; Mills & Tissot, 1995; Stephens et al., 1999). These studies were conducted on the MAT, not the revised NNAT. Perhaps the test revision affected the correlation, which may have influenced cultural fairness in the NNAT.

Also, it is possible that the correlation between these tests is due to the fact that all are measuring similar components of intelligence leading to some overlap between the tests, while at the same time measuring some unknown variables that lead to greater cultural fairness.

The majority of past research and statements in the administration manuals appear to indicate that the Raven’s Standard Progressive Matrices and the Naglieri Nonverbal Abilities Test may be effective instruments in selecting for ethnically diverse students who may be gifted (Karnes & Whorton, 1988; Mills et al., 1993; Mills & Tissot, 1995; Naglieri, 1996; Prewett, 1995; Stephens et al., 1999). However, research is lacking investigating whether either of these nonverbal abilities tests is more effective in selecting for ethnically diverse gifted students than a more traditional measure of achievement, such as the Iowa Test of Basic Skills.

Therefore, the purpose of this study was to compare the effectiveness of the Raven’s Standard Progressive Matrices, the Naglieri Nonverbal Abilities Test, and the Iowa Test of Basic Skills in selecting for ethnically diverse students who may be gifted.

Method

Sample

The participants in the present study were 175 students enrolled in Grades 3-5 and Grade 8 in a Midwestern school district serving a small city of approximately 40,000 and surrounding rural areas. The 6th and 7th Grades were not included in this study because the school district did not administer group achievement tests at these grade levels.

This school district was chosen because of its large Hispanic population (40%). The elementary school and middle school that participated in the study served the highest concentration of Hispanic and low-income students in the district. Ethnicity was divided into three categories: Caucasian (n=102), Hispanic (n=70), and Other (n=5). Students were determined to be of Caucasian or Hispanic origin as defined by parents or self-report.

The category of Other was created by Lewis (1999) to include all participants who did not fall into either of the other two groups. Because only five students fell under the heading of Other, it was combined with the Hispanic category to form the group Ethnically Diverse. Complete data for two of these students was unavailable reducing this category to n=73.
Research

Procedure

Archival data on the Raven’s, the NNAT, and each student’s ethnicity was collected from an earlier study by Lewis (1999), along with archival scores from the Iowa Test of Basic Skills administered by the school district. Lewis and her co-researcher administered the Raven’s Standard Progressive Matrices, Naglieri Nonverbal Abilities Test, and the Culture Fair Intelligence Test to participants in Grades 3-8. Data from the CFTT were not used in this study.

The order of administration was counterbalanced. The tests were administered to students during regular class time. An attempt was made to include all students in Grades 3-8; however, a few students did not participate in all of the tests due to illness or behavioral difficulties. The administration of each test took less than one hour for every grade. The school district provided a computer printout detailing each participant’s race as reported by parents for grades 3-5 and self-report was used for grades 6-8.

The district provided a printout of Iowa Test of Basic Skills scores for each of the participants. Once these two files were combined for Grades 3-5, and 8, the students’ names were removed to ensure confidentiality. Only the Composite scores were used in the analyses.

Students’ scores on the Raven’s, the NNAT, and the ITBS were compared to determine which assessment identified the greatest number of students in each cultural category at or above the 80th percentile level, this being the inclusive level used by Stephens et al. (1999).

Instrumentation

The Raven’s Standard Progressive Matrices is a nonverbal untimed test that measures higher-level thinking skills (Raven et al., 1998). The test has 60 items divided into five Sets (A, B, C, D, and E). Each Set consists of 12 items, arranged in a progressively more difficult order. The Raven’s is group administered and relatively easy to score. United States norms are reported for ages 6.5 through 16.5 for the Raven’s; however, these norms are approximated from various local norms rather than being derived through a systematic norming process.

The norming process of the Raven’s has been cited as a major limitation of the instrument (Mills et al., 1993). The Raven’s has shown acceptable reliability across geographic boundaries and test-retest reliability coefficients ranging from .76 to .91, with the highest values being found for older groups (Raven et al., 1998).

The Naglieri Nonverbal Ability Test is “a brief, culture-fair, nonverbal measure of school ability” (Naglieri, 1996, p.3) that is group-administered and timed. The NNAT is divided into seven levels: Level A for Kindergarten, Level B for Grade 1, Level C for Grade 2, Level D for Grades 3-4, Level E for Grades 5-6, Level F for Grades 7, 8, 9, and Level G for Grades 10, 11, 12. Each level has 38 items. For this study archival data were only available for Levels D and F. Grades 3-5 were administered Level D, while Grade 8 was administered Level F. Directions for Administering and Scoring the Naglieri Nonverbal Abilities Test provides for these scoring procedures (Naglieri, 1996).

The NNAT was normed using a nationally stratified, random sampling scheme. Reliability was computed in the form of internal consistency estimates at each level and calculated separately for age and grade. Internal consistency estimates for total scores range from .80 to .93 (Naglieri, 1996). No test-retest reliability coefficients are presented in the administration manual, an omission that is considered a weakness of the instrument (Stinnett, 1997).

The Iowa Test of Basic Skills (ITBS) is a group administered test of academic achievement that consists of three forms (K, L, M). The purpose of the ITBS is “to provide a comprehensive assessment of student progress in the basic skills” (Hoover, Hieronymous, Frisbie, & Dunbar, 1996, p. 2). The ITBS is available for Levels 5-14, which roughly correspond to the ages of the children to whom the test may be administered. The Complete Battery consists of the following sections: Listening, Word Analysis (Levels 5-8), Vocabulary, Reading, Language, Mathematics, Social Studies, Science and Sources of Information (Levels 7-14).

In addition, sections on Listening Assessment and Writing Assessment are included for Levels 9-14. Subtest scores can be combined to yield a Reading Total score, a Language Total score, a Mathematics Total score, and a Total Composite score. Raw scores obtained from the Iowa Test of Basic Skills are converted into percentages that can be norm referenced. The ITBS was normed using a nationally stratified sample representative of the general population. Most test-retest reliabilities for the subtests fall between .80 and .90 (Hoover et. al, 1996).

Results

To determine the effectiveness of the Raven’s Standard Progressive Matrices, the Naglieri Nonverbal Abilities Test, and the Iowa Test of Basic Skills in selecting for ethnically diverse children that may be gifted, a Cochran Q analysis was performed on the data (see Table 1). Results of the Cochran Q analysis indicated a significant difference among tests in identifying ethnically diverse students who may be gifted, $X^2(2, N=73)=17.29, p<.001$. A similar Q test was also performed on the Caucasian children’s data (see Table 1).

These results also indicated a significant difference among tests in identifying Caucasian children who may be gifted, $X^2(2, N=102)=17.44, p<.001$. While the Raven’s identified 54% more gifted Caucasian children than Ethnically Diverse children, the NNAT identified 262% and the ITBS identified 370% more Caucasian children. Furthermore, the Raven’s identified 560% more potentially gifted children from diverse backgrounds than either NNAT or the ITBS composite.

The Pearson Product Moment correlation coefficients among the three tests are displayed in Table 2. The correlation between the Raven’s and the ITBS was less than the correlation between the NNAT and the ITBS, indicating that the latter two had more in common. Note that both the ITBS and the NNAT could identify far lower percentages of potentially gifted children compared with the Raven’s in both samples employed in the present study (see Table 1).

A two-way analysis of variance with repeated measures on one factor was also utilized to compare the two groups of children (Caucasian vs. Ethnically Diverse) on their percentile scores of the three tests (Raven’s vs. NNAT vs. ITBS). The two groups serve as the independent factor while the three tests serve as the repeated factor. Results of the analysis indicated a significant “group” main effect, $F(1, 173)=15.54, p<.001$; a significant “test” main effect, $F(2, 346)=62.65, p<.001$; and a significant “group x test” interaction effect, $F(2, 346)=3.25, p<.05$ (see Table 3 and Figure 1).

Because the interaction was significant ($p<.05$), “simple effect tests” were performed comparing the Caucasian and Ethnically Diverse children on each of the three tests. Results indicated that the two groups of children differed significantly on the Raven’s, $F(1, 362)=3.88, p<.05$; differed significantly on the NNAT, $F(1, 362)=8.67,
Research

p < .01; and differed significantly on the ITBS, F(1, 362)=20.95, p<.001. Figure 1 demonstrates that the difference between the two groups was most pronounced on the ITBS test.

Discussion

This study compared the ability of three different group tests to screen for potentially gifted ethnically diverse children. Two tests were measures of nonverbal intelligence (Raven Standard Progressive Matrices and Naglieri Nonverbal Abilities Test) and the third test was an achievement test that is given annually to tens of thousands of school children across the country (Iowa Test of Basic Skills).

Group achievement tests such as the ITBS are one means of gathering data on students' ability to function at the high end of academic performance and thus may indicate the need for more advanced instruction. Ethnically diverse students may not score well on such verbal measures. In order to meet best practice guidelines, additional methods need to be used to provide multiple ways that all students with the potential to benefit from gifted programming might be observed (Johnsen, 2004).

In practice, high percentile scores from district-wide achievement tests are often used to screen for, and sometimes even place children in gifted programs, with or without corroborating assessments. The results of this study indicate that the Raven's, the NNAT, and the ITBS are not equally effective in selecting potentially gifted children. Not surprisingly, the ITBS identified the lowest percentage of potentially gifted children, possibly indicating problems with the process of teaching and learning rather than intelligence since the ITBS does not purport to assess intelligence.

Although both the Raven's and the NNAT are measures of nonverbal intelligence that use different kinds of geometric designs, in the present study the Raven's selected more ethnically diverse students with potential to be successful in gifted programs than the NNAT, despite the findings of Naglieri and Ford (2003) that showed potentially gifted Caucasian, Hispanic, and African-American children were discovered in similar proportions. The progressive nature of the Raven's, in which students can learn from previous items, may have captured the essence of gifted thought processes more effectively than the structure of the NNAT.

Results of this study indicated that the

<table>
<thead>
<tr>
<th>Table 1: Cochran Q Test Comparing the Percentages of Caucasian and Ethnically Diverse Students Scoring in the 80th Percentile and Above on the Raven's, NNAT, and ITBS Composite</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Group</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Caucasian (N=102)</td>
</tr>
<tr>
<td>Ethnically Diverse (N=73)</td>
</tr>
</tbody>
</table>

*** p = < .001

<table>
<thead>
<tr>
<th>Table 2: Correlation among the Raven's, NNAT, and ITBS Composite Percentile Scores (N = 175)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Test</strong></td>
</tr>
<tr>
<td>Raven's</td>
</tr>
<tr>
<td>ITBS Composite</td>
</tr>
</tbody>
</table>

*** p < .001

<table>
<thead>
<tr>
<th>Table 3: Mean Percentile Scores on the Raven's, NNAT, and ITBS Composite Comparing Caucasian and Ethnically Diverse Children</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Group</strong></td>
</tr>
<tr>
<td><strong>M (SD)</strong></td>
</tr>
<tr>
<td>Caucasian (N= 102)</td>
</tr>
<tr>
<td>Ethnically Diverse (N= 73)</td>
</tr>
</tbody>
</table>

| Figure 1: Comparing Caucasian and Ethnically Diverse Children on the Raven's, NNAT, and ITBS Composite Percentile Scores |

![Figure 1: Comparing Caucasian and Ethnically Diverse Children on the Raven's, NNAT, and ITBS Composite Percentile Scores](image-url)
Raven's Standard Progressive Matrices was a more effective means of selecting for ethnically diverse children who may be gifted than one example of a traditional achievement test or even the newer Naglieri Nonverbal Abilities Test. The higher level thinking skills demonstrated on this assessment suggest that these students may benefit from increased support and placement in gifted programs.

It is recommended that the Raven Standard Progressive Matrices be considered as one of the methods employed by a district to select for children who would benefit from gifted programming. While the Raven's is not bias free, the proportion of ethnically diverse students was greater for the same sample used in this study than the other two tests. By testing entire classes, as is the procedure with the traditional achievement tests, all the students in those classes have the opportunity to demonstrate their proficiency on the particular skills assessed by the Raven's as is recommended by Johnsen (2004).

References


