JSGE
The Journal of Secondary Gifted Education
Vol. XV, No. 2, Winter 2004, pp. 61–69.

DISCOVER in Middle School:

Identifying Gifted Minority Students

Ketty M. Sarouphim
Lebanese American University

The purpose of this study was to examine the validity of the grades 6–8 version of DISCOVER, a performance-based assessment, and investigate its effectiveness in identifying gifted minority students. Questions examined the alignment between DISCOVER and Gardner's (1983) theory of multiple intelligences (MI) and assessed gender and ethnic differences. The sample consisted of 395 middle school students, predominantly Mexican Americans and Native Americans belonging to lower socioeconomic classes, from schools in Arizona. Results supported a good fit between DISCOVER and MI theory. That is, students identified as gifted in one intelligence were not necessarily identified in other intelligences. The 2 x 3 MANOVA (gender by ethnicity) showed no significant interaction or main effect for ethnicity. However, a main effect for gender was found with males outperforming females in math. No overall gender or ethnic differences in identification were revealed. In total, 12.4% of the participants were identified, suggesting that using DISCOVER might diminish the long-standing problem of minority underrepresentation in gifted programs.

Educators have long acknowledged the problem of minority students' underrepresentation in programs for the gifted (Baker, 1996; Bernal, 2002; Cummings, 1991; Ford & Harmon, 2001; Maker, 1996). The problem is significant, especially since underrepresentation has increased over the years with respect to certain ethnic groups such as Blacks, Hispanics, and Native Americans (Ford, Harris, Tyson, & Trotman, 2002). Some educators estimate that students from these culturally diverse groups are underrepresented by as much as 30–70% relative to their percentage in the population (Gabelko & Sosniak, 2002).

The purpose of this study was to investigate the effectiveness of DISCOVER, a performance-based assessment, in identifying gifted minority middle school students and reducing their underrepresentation in programs for the gifted. DISCOVER is an acronym that stands for Discovering Intellectual Strengths and Capabilities through Observation while allowing for Varied Ethnic Responses (Maker, Nielson, & Rogers, 1994).

Reasons for Underrepresentation

Ford et al. (2002) argued that the main reason for minority underrepresentation is a "deficit perspective" that has influenced directly or indirectly the access of culturally diverse students to gifted programs. Educators who hold this perspective assume that students from diverse and economically disadvantaged populations are cognitively inferior because many score low on standardized tests and fail to meet the traditional criteria for placement in gifted programs, that is, scoring on the 97th percentile or above. One symptom of the cognitive deficit hypothesis is the narrow definition of intelligence and giftedness adopted in most schools. A major limitation of these definitions is that they do not take into consideration cultural factors in conceiving and demonstrating superior cognitive abilities (Bernal, 2002). As a result of the widespread belief in these traditional conceptions, identification procedures in most school districts (about 90%) still rely heavily on the scores of standardized tests, a practice that limits the access of culturally diverse students to programs for the gifted and keeps the demographics of these programs mostly White (Ford & Harmon, 2001; Ford et al.).

Alternative Assessments for Gifted Identification

Many educators have advocated replacing standardized tests with a new breed of instruments, namely authentic assessment, also called alternative and performance-based assessment (Ford & Harmon, 2001; Maker, 1996, Sarouphim, 1999). The rise of authentic assessment (e.g., Borland & Wright, 1994; Clasen, Middelton, & Connell, 1994; Hafenstein & Tucker, 1994; Maker, 1992; Reid, Udall, Romanoff, & Algozzine, 1999) has coincided with the emergence of alternative theories of intelligence, such as Gardner's (1983) theory of multiple intelligences (MI) and Sternberg's triarchic theory of intelligence (1986). Along the same lines, recent conceptions of giftedness have also played a major role in promoting the development of authentic assessment. For example, Maker (1993) defined giftedness as "the ability to solve complex problems in effective, efficient, elegant, and economical ways" (p. 71).

Research on alternative assessment has revealed that, more often than not, minority students fare better on these measures than on traditional standardized tests (Borland & Wright, 1994; Clasen et al., 1994; Reid et al., 1999; Sarouphim, 2001), a factor that earned authentic assessment the reputation of being bias-free and expanded its use among minority groups. However, authentic assessment is not without its drawbacks. Among the many criticisms of these measures are their high cost, domain underrepresentation, lack of sound psychometric properties, and their long and laborious administration (Dunbar, Koretz, & Hoover, 1991; Frechtling, 1991).

Evidence Against the Use of Alternative Assessments

Research on the effectiveness of alternative assessment has yielded mixed results, a factor that has led some researchers to warn against the use of these measures. For example, Plucker, Callahan, and Tomchin (1996) did not find evidence supporting the sound psychometric properties of a battery of alternative instruments based on the theory of multiple intelligences. Participants (N = 1,813) were assessed using the Multiple Intelligence Assessment Technique, based on the work of Project Spectrum and local modifications of the DISCOVER assessment. Student performance was rated as "not evident or not observed," "evident," or "extremely evident." The results showed high internal consistency for the scales' scores, and factor analysis confirmed the presence of the linguistic and logical-mathematical subscales, but the presence of the spatial and interpersonal scales were not supported. Generally, the correlations between the different subscales were sufficiently low to provide evidence for discriminant validity, but a relatively high correlation was found between the ITBS language subscale and the math performance assessment, as well as between the math checklist and the linguistic checklist. No significant gender or ethnic differences were found.

The study raised questions about the structural and construct validity of MI-based assessments, pointing to the methodological difficulties of assessing the psychometric properties of such instruments. The researchers concluded that although "MI theory and alternative assessments may hold substantive implications for education of gifted students . . . educators using MI theory, alternative assessments, and combinations of the two should subject the programs to rigorous evaluation" (Plucker, Callahan, & Tomchin, 1996, p. 87).

Evidence Supporting the Use of Alternative Assessment

In a study of whether alternative assessment might diminish the problem of underrepresentation of minority students in programs for the gifted, Reid et al. (1999) compared a traditional measure, the Matrix Analogies Test–Short Form (MAT-SF), with an alternate assessment, the Problem Solving Assessment (PSA), in identifying culturally and linguistically gifted students. The PSA is a measure grounded in Gardner's theory of multiple intelligences and the conceptual framework of the DISCOVER assessment. It consists of tasks designed to measure linguistic, spatial, and logical-mathematical intelligences. To be recommended for placement in a gifted program through the use of PSA, students need a rating of "always evident" or "strongly evident" problem-solving behaviors in two or more of the intelligences assessed. A stanine performance of 9 in the MAT-SF was used as the criterion for identification.

The sample consisted of 600 students of White, Black, Hispanic, American Indian, and Asian origin. The results showed that, through the use of MAT-SF, only 22% of the students met the criteria for identification, whereas about half of the students were identified as gifted through the use of the PSA. However, such a high percentage of identified students is questionable and does not in itself validate the assessment. Also, significant differences were found in the distribution of identified students. Using the MAT-SF, 11% of identified students were minority, whereas the PSA recommended 39% of minority students for placement in gifted programs. However, the study has an inherent weakness: the lack of data on the predictive validity of the PSA. Also, a similarly high percentage of identified students using the MAT-SF might result by lowering the cut-score used. Despite these limitations, the researchers concluded that available anecdotal evidence on the predictive validity of the PSA was favorable: Most students who were placed in programs for the gifted using the PSA were successful.

In another study, Hafenstein and Tucker (1994) assessed the effectiveness of a nontraditional assessment for measuring multiple intelligences. Trained observers assessed 3-, 4-, and 5-year-old children as they worked on tasks in the seven intelligences. Collected data included interviews with teachers and reports from the students' parents. Following the assessment, observers classified children's abilities as "not evident," "evident," and "extremely evident" (i.e., gifted). In mid-year, teachers were asked to rate the children using the same classification. Mean differences between the teachers' two ratings were minimal; the first mean was 3.67 and the second was 3.74.

Also, content analysis revealed a great similarity between observers' and teachers' ratings, as well as between the parents' and teachers' reports, suggesting that the use of the assessment led to adequate placement of children. Interviews with parents indicated that they believed that their children were well placed and thriving in the gifted program. Interviews with teachers showed similar results. Regression analysis suggested that the beginning-of-year assessment was predictive of future performance, $F(1,34) = 4.951$, $p < .03$. The researchers concluded that the performance-based assessment used in this study was an effective process for identifying young gifted children.

Research on DISCOVER

The alternative assessment, DISCOVER, which is the focus of this study, is grounded in MI theory and based on Maker's (1993) definition of giftedness. The assessment was designed to reflect Gardner's (1983) description of the "core capabilities" of each intelligence. DISCOVER was developed to identify gifted students from culturally diverse groups. Since its inception, DISCOVER has been administered to thousands of students from diverse populations. The data collected have served as the basis for research on the reliability and validity of the instrument.

A series of studies were conducted to examine the psychometric properties of the K–2, 3–5, and 9–12 versions of DISCOVER. These studies focused on the alignment between DISCOVER and MI theory, gender differences, ethnic differences, and concurrent validity of DISCOVER with the Raven's Progressive Matrices. Even though other research on DISCOVER exists, this section will be limited to the review of studies on the aforementioned questions because their focus was similar to that of the present investigation (for a thorough review of research on DISCOVER, see Sarouphim, 2002). However, one additional study on the interrater reliability of DISCOVER is worth mentioning here. Griffiths (1996) compared the ratings observers gave to students on the spatial activities and those marked by independent raters who watched videotapes of the recorded administration. The results showed high interrater agreement, ranging from 80% to 100%, with the highest agreement found between the observers and independent raters with the most expertise in the administration of DISCOVER.

Fit between DISCOVER and MI theory. Sarouphim (2000) investigated the alignment of DISCOVER with the theory of multiple intelligences through a series of interobserver correlations. The sample consisted of 254 elementary students, predominantly from economically disadvantaged Native American and Hispanic groups. All participants took either the K–2 or the 3–5 version of DISCOVER, depending on their grade level. The results showed low interobserver correlations across grade levels between the activities that measured different intelligences (e.g., linguistic and spatial activities) and moderate to high correlations between activities that measured related intelligences (e.g., oral linguistic and written linguistic), indicating that students who were identified as gifted in one intelligence were not necessarily identified as gifted in the other intelligences. The results suggested that the different DISCOVER activities with distinguishable cognitive tasks may measure different intelligences, a finding that might provide support to the consistency between DISCOVER and Gardner's MI theory.

Concurrent validity. In another study, Sarouphim (2001) examined the concurrent validity of DISCOVER with the Raven's Progressive Matrices. The study also examined gender differences in identification rates and the percentage of minority students identified by DISCOVER and traditional standardized tests. The results, which were based on a sample of Native American and Hispanic students, showed a high correlation between the students' scores on the Raven's and their ratings in the spatial activities of DISCOVER and low correlations between the students' Raven scores and their ratings in the linguistic activities of DISCOVER, providing evidence for the convergent and discriminant validity of DISCOVER. Using the criterion of a "definitely" rating in at least two of the activities, the results also showed that, through the use of the DISCOVER assessment, 22.9% of the students were identified. In addition, no significant gender differences were found in identification, possibly indicating that the assessment is mostly fair and does not discriminate against gender or ethnicity.

Gender and ethnic differences. Finally, Sarouphim (2002) investigated the effectiveness of the 9–12 version of DISCOVER. The sample consisted of 303 ninth graders, predominantly Hispanic and Native American students. The results provided evidence for an alignment of the assessment with the theory of multiple intelligences. Also, no overall gender or ethnic differences were found in identification. In addition, the results suggested that the use of the DISCOVER assessment might help in reducing the problem of minority students' underrepresentation in programs for the gifted, as 29.3% of the high school students who participated in the study were identified as gifted.

The Current Study

The purpose of this study was to examine the validity of the 6–8 version of DISCOVER and investigate the effectiveness of this assessment in identifying gifted middle school students from culturally diverse groups. Three questions guided this inquiry: (1) Are DISCOVER and MI theory aligned? That is, does DISCOVER tap into the different intelligences as identified by Gardner? (2) Do gender differences appear through the use of DISCOVER? (3) Do ethnic differences appear through the use of DISCOVER?

Method

Participants

The sample consisted of 395 male (54%, *n* = 214) and female (46%, *n* = 81) sixth, seventh, and eighth graders from 18 schools in Arizona. Participants were mostly of low socioeconomic status, as evidenced by their place of residence and their participation in their school's free lunch program (85%). Participants were Native Americans from the Navajo tribe (47%, *n* = 186) attending 8 schools in northern Arizona and Mexican Americans (38.5%, *n* = 151) and White Americans (14.5%, *n* = 58) attending 10 schools in southern Arizona. The mean ages were 13.1, 13.3, and 13.6 years for White, Mexican American, and Native American students, respectively. The mean age of females was 13.1 years and 13.2 years for males.

Instrument

DISCOVER is a performance-based assessment consisting of five activities designed to measure individuals' problem-solving abilities in each of the following intelligences: Pablo® (spatial), Tangrams (spatial/logical-mathematical), Math (logical- mathematical), Storytelling (oral linguistic), and Storywriting (written linguistic). Each activity proceeds through a series of tasks that progress from structured to more open or "fuzzy" problems. To avoid observer bias, observers rotate at the completion of each activity so that each student is assessed only once (i.e., during one activity only) by the same observer. The following is a brief description of each activity (for a thorough description of DISCOVER, see Maker, 1992, and Sarouphim, 1999).

Pablo. The material for this activity consists of colored cardboard pieces of different shapes, designs, and sizes. Students are asked to make different constructions (e.g., geometrical designs, a container, a machine, and a construction of their choice) using the Pablo pieces.

Tangrams. Each student is given a set of Chinese Tangrams (21 pieces of three different shapes: triangles of three different sizes, squares, and parallelograms). Students are requested to make a parallelogram using as many Tangram pieces as possible. Then, each student is given a booklet of six puzzle sheets arranged in ascending order of difficulty and asked to solve them. Students who complete this task are then provided with "Challenge Sheets," which consist of more difficult puzzle problems.

Storytelling. Students are given an array of toys and are asked to describe one and then two of their toys using as many descriptors as possible. Then, students are asked to tell a story of their choice that incorporates some or all of the toys they have been given.

Storywriting. Students are asked to produce a written piece of their choice (story, poem, etc.) about a topic of their choice.

Math. Worksheets consisting mostly of open-ended numerical problems are used to assess this intelligence. The problems increase in openness and difficulty, with the last problem consisting of creating as many problems as possible (fractions, multiplication, division, etc.) having a prespecified number as the answer.

Procedures

All participants were assessed through DISCOVER for identification purposes. Trained observers administered the spatial, logical-mathematical, and verbal linguistic activities in class and took notes while students worked in groups of four or five students, with a ratio of 1:4 or 1:5 (one observer to four or five students). All observers completed a 2-day training during which they learned about the conceptual and applied aspects of the assessment, followed by at least four in vivo sessions during which they "shadowed" expert observers and watched them in action. Observers were all holders of doctoral degrees in education or graduate students pursuing advanced degrees in education. In most classes, teachers read the instructions, but, in a few instances where teachers preferred not to take part in the process, one of the DISCOVER observers gave the instructions. The possible source of error that might result from an observer rather than the classroom teacher giving instructions is kept minimal since the DISCOVER administration procedures are standardized. For

Table 1

Observers' Interrating Correlations

Activity	1	2	3	4	5
Pablo	–	0.239**	0.267**	0.074	0.012
Tangrams	–	–	0.102	0.084	0.083
Math	–	–	–	0.183*	0.023
Story	–	–	–	–	0.258**
Writing	–	–	–	–	–

Note. *p < 0.05. **p < 0.01

accuracy purposes, the administration sessions were videotaped and the students' stories were tape-recorded and transcribed verbatim at a later time. A day preceding or following the group activities, the students worked individually on the written linguistic and math worksheet components of DISCOVER.

Following data collection, the DISCOVER observers met and discussed the students' performances. They then ranked students' performances in each of the activities according to a 4-category rating scale of "unknown," "maybe," "probably," and "definitely," with the last rating category being the highest and corresponding to superior problem-solving ability or giftedness. Usually, students given the "definitely" rating category in at least two of the activities are identified as gifted; however, the identification criteria are flexible to a certain extent (e.g., in some schools, students given three "definitely" ratings are identified as gifted) and depend on the school district identification procedures, as well as the breadth of programs for the gifted offered at each particular school (Sarouphim, 2001). However, in this study, the criterion for identification used was set at a rating of "definitely" in at least two of the activities. Data were coded as follows: 1 for "unknown," 2 for "maybe," 3 for "probably," and 4 for "definitely."

To assign a rating, observers are guided by a checklist they complete for each child. Items on the checklist represent superior problem-solving behaviors (process) and characteristics of products. For example, in Pablo, observers note how the final construction was produced and whether the constructions are three-dimensional, complex, and original and whether they incorporate many pieces. In Tangrams, observers note the number of puzzle sheets solved, the strategies used (e.g., fitting pieces without physically rotating them), the time it takes students to solve them, and the number of Tangram pieces used to complete a square or a triangle. In Storytelling and Storywriting, observers look for fluency, well-constructed plots, appropriate sequence of events, and the quality of words and sentences. In Math, strategies (e.g., creating problems and solving them), as well as the number of problems solved, are taken into consideration. Even though the checklist constitutes an important element in ranking students, it is not considered exclusive in the decision-making process. That is, the observers' classification is also based on noting superior problem-solving behavior not included in the checklist. The rationale is that giftedness is not limited to a few behaviors, so observers need to keep an open mind and note all performance that might denote superior ability.

Results

To assess the alignment of DISCOVER with MI theory, the ratings given to students by trained observers in the five DISCOVER activities were correlated, yielding a matrix of observers' interrating correlations. For gender and ethnic differences in activity (i.e., per intelligence), a 2 x 3 MANOVA (gender by ethnicity) and follow-up ANOVA were calculated, whereas for gender and ethnic differences in identification (i.e., the number of identified students per gender and ethnicity), the chi-square statistic was used.

Alignment of DISCOVER With MI Theory

As seen in Table 1, all correlations were low, though some were significant. The highest correlation was found between the two activities of Math and Pablo, $r = .267$, $p < .01$, and the lowest between the two activities of Math and Storywriting, $r = .023$, ns. A significant correlation was found between the two activities of Storytelling and Storywriting, $r = .258$, $p < .01$, both measuring linguistic intelligence. The results suggest that students' ratings (i.e., performance) were comparable in the activities that assessed a similar intelligence and discrepant in the activities that assessed different intelligences.

Gender and Ethnic Differences by Activity

Table 2 shows the students' mean ratings and standard deviations in all activities. The means ranged between "maybe" and "probably," with White males obtaining the highest mean ratings in the Math activity and Native American males obtaining the lowest mean ratings in the Storywriting activity. The multivariate analysis of variance showed no significant interaction effect between gender and ethnicity, $F(6,562) = 1.61$, $p > .05$. Similarly, the results revealed no significant main effect for ethnicity, $F(6, 562) = 1.03$, $p > .05$; however, a significant main effect for gender was found, $F(4,541) = 9.67$, $p < .01$. Post-hoc univariate analysis showed that males outperformed females in Math, $F(1,373) = 7.57$, $p < .01$, yielding an effect size of .69.

Table 2

Students' Mean Ratings and Standard Deviations by Gender and Ethnicity

	Mexican American	Native American	White American	All
Pablo				
M	2.65 (0.97)	3.01 (0.84)	2.90 (0.73)	2.85 (0.84)
F	2.71 (0.91)	2.79 (0.89)	2.84 (0.80)	2.78 (0.86)
All	2.68 (0.94)	2.90 (0.86)	2.87 (0.76)	2.81 (0.85)
Tangrams				
M	2.65 (0.86)	2.58 (0.72)	2.70 (0.91)	2.64 (0.83)
F	2.71 (0.91)	2.60 (0.81)	2.65 (0.79)	2.65 (0.83)
All	2.68 (0.88)	2.59 (0.76)	2.67 (0.85)	2.64 (0.83)
Math				
M	2.97 (0.65)	3.01 (0.98)	3.11 (1.02)	3.03 (0.88)
F	2.54 (0.88)	2.48 (0.79)	2.63 (1.13)	2.55 (0.93)
All	2.75 (0.76)	2.74 (0.88)	2.87 (1.07)	2.79 (0.90)
Storytelling				
M	2.92 (0.87)	2.85 (0.71)	3.05 (0.87)	2.94 (0.81)
F	2.98 (0.93)	2.71 (0.89)	3.02 (0.92)	2.90 (0.91)
All	2.95 (0.90)	2.78 (0.80)	3.03 (0.89)	2.92 (0.86)
Storywriting				
M	2.51 (0.76)	2.36 (0.84)	2.43 (0.73)	2.43 (0.77)
F	2.74 (0.82)	2.64 (0.88)	2.86 (0.81)	2.74 (0.83)
All	2.62 (0.79)	2.50 (0.86)	2.64 (0.77)	2.58 (0.80)

Note. Unknown = 1, Maybe = 2, Probably = 3, Definitely = 4.

Gender and Ethnic Differences by Identification

The number of males and females identified as gifted (i.e., given at least two "definitely" ratings) in the DISCOVER activities was calculated. The results revealed that the number of identified males was not significantly higher than the number of identified females (see Table 3). The Mexican American group had the highest percentage of identified students, followed by Native Americans and Whites. The chi-square test showed no significant overall gender or ethnic differences in identification, $X^2(2,94) = .346$, ns. In total, 12.4% of participants were identified as gifted.

Discussion

The purpose of this study was to examine the validity of the grades 6–8 version of DISCOVER, a performance-based assessment, and investigate its effectiveness in identifying gifted minority students. The research questions focused on the alignment between DISCOVER and MI theory and an examination of gender and ethnic differences. The results showed a good fit between DISCOVER and MI theory and an absence of significant gender and ethnic differences in identification. However, the univariate analysis showed that males outperformed females in the Math activity of DISCOVER. Finally, the percentage of students identified as gifted through the use of DISCOVER was higher than the traditional 3% designated through the use of traditional standardized tests.

In this study, the low observers' interrating correlations indicated that a student who was given a high rating in one intelligence (i.e., identified as gifted) was not necessarily given the same high rating in the other intelligences. In other words, the findings reveal that DISCOVER might tap into a variety of intelligences, suggesting a good fit between DISCOVER and MI theory. On the other hand, the finding that all correlations

Table 3

Gifted Participants by Gender and Ethnicity

	Mexican American		Native American		White American		Total	
	All	Gifted	All	Gifted	All	Gifted	All	Gifted
Males	78	11	101	13	35	4	214	28
Females	73	9	85	10	23	2	181	21
Total	151	20	186	23	58	6	395	49

were low, including those found significant, is problematic. According to Gardner (1983), some of the intelligences are tightly related, such as spatial and logical-mathematical, which he called "twin" intelligences, so the correlations in ratings given to students in Pablo, Tangrams, and Math were expected to yield moderately high values, rather than the relatively low values found in this study. The same reasoning applies to the activities of Storytelling and Storywriting, which both measure linguistic intelligence. Also, statistical significance of the correlations in this study might be due to the relatively large sample size, rather than a true correlation in the population. In further studies with similar analyses, further exploration of this issue is needed to explain this problematic finding.

Moreover, the activities included in DISCOVER do not cover the whole spectrum of Gardner's recognized intelligences, namely, bodily-kinesthetic, musical, naturalist, existentialist, and the personal intelligences. The rationale for focusing on spatial, mathematical, and linguistic intelligences is that a good match must exist between an assessment and a placement program and, at this time, programs for the gifted emphasize mostly these three intelligences. However, one recommendation that stems from this study is that, for DISCOVER to be faithful to MI theory, all the intelligences must be assessed. Therefore, activities for all intelligences need to be developed. Similarly, another recommendation is that school officials who intend to use DISCOVER for placement purposes need to be made aware of the limitations of the assessment, especially if a discrepancy exists between the focus of the gifted program and that of DISCOVER.

An interesting finding was the absence of gender and ethnic differences in identification. Educators have long deplored the ethnic bias in identification (Bernal, 2002; Ford & Harmon, 2001). The absence of such a bias in this study, as well as the high percentage of identified students, suggests that the use of DISCOVER might help in reducing the problem of minority underrepresentation in programs for the gifted. This finding is compatible with the results of other studies in which performance-based assessments were used for identification purposes (Borland & Wright, 1994; Clasen et al., 1994; Hafenstein & Tucker, 1994; Reid et al., 1999). In these studies, the final pool of identified students was larger than that usually found through the administration of standardized tests.

However, the identification of a large number of minority students does not justify in and of itself the use of DISCOVER or any other performance-based assessment. Rather, the use of an instrument must be justified by evidence of its effectiveness. Even though preliminary research results provide evidence for the high reliability and validity of DISCOVER, data on the predictive validity of the assessment, which would be a better indicator of its effectiveness, are still missing. At this time, a longitudinal study is underway to follow the academic progress of a group of Native American students (the "Step Up" group) who were all identified through DISCOVER and placed in different programs for the gifted. But, until the results are out, solid conclusions on the predictive validity of DISCOVER cannot be drawn.

A noteworthy finding is that White students performed as well as minority participants in the DISCOVER activities, which might indicate that DISCOVER could be used for the identification of majority students, as well. However, the sample of White students in this study was too small to warrant such an assumption. In further studies, larger numbers of White students must be included in the research sample to clarify this issue.

The finding that males received higher ratings in the Math activity is congruent with a body of research on gender differences in mathematical intelligence (e.g., Lubinski & Benbow, 1992). In previous studies on the K–2, 3–5, and 9–12 versions of DISCOVER, gender differences in the Math activity were not found (Sarouphim, 1999, 2002). One explanation might be that the Math activity in the middle school version of DISCOVER is more sensitive to students' strengths and taps better into this intelligence. Such an explanation entails a revision of the Math activity in the other versions of DISCOVER because, if a difference truly exists in the population, an assessment that shows such a difference has more valid scores than one that does not. However, other explanations specific to the population of middle school students might be considered. For example, during adolescence, girls start showing interest in boys and their behavior becomes gender-specific. That is, they are more careful to abide by behaviors labeled "feminine" in the culture (Lynch, 1991). Traditionally, superiority in math has been associated with males (Eccles, 1993), and, consequently, an outstanding performance in math might not be considered "feminine," which might explain the drop in the girls' math performance in the 6–8 version of DISCOVER. However, further research is needed to explain such a finding.

Moreover, in previous research on DISCOVER, the percentages of identified students were higher (i.e., 22.9% in K–5 and 29.3% in 9–12). An explanation of this finding might be grounded in the population studied. Middle school is the bridge between elementary and secondary school and constitutes a period of academic transition (Coleman, 2001). Researchers have long noticed a drop in the academic performance of middle school students. Some have attributed this decline to a traditional classroom environment that does not match early adolescents' needs and interest (Eccles, Wigfield, Midgley, Reuman, MacIver, & Feldlaufer, 1992). Therefore, the lower percentage of identified students in this study might reflect the general academic decline that accompanies the transition from elementary to middle school. However, the reason might also be related to the specific tasks of the 6–8 version of DISCOVER. Again, further examination of the activities in DISCOVER through a thorough content analysis of the tasks incorporated is needed to clarify this point.

The results of this study provide positive evidence for the use of the DISCOVER assessment with culturally diverse groups, namely Mexican Americans and Native Americans. The implications are of significance to practitioners and researchers alike who have been striving relentlessly to find means for rectifying the long-standing injustice of minority underrepresentation in programs for the gifted. Thus, DISCOVER can be used for identification purposes, especially for the placement of students in gifted programs that match the theoretical bases of the assessment (i.e., multiple intelligences). A paradigm shift in assessment procedures can contribute significantly to the solution of minority students' underrepresentation in gifted programs. In the continuous struggle to establish equity in gifted education, the use of authentic assessment seems to be promising. Instruments such as DISCOVER can greatly contribute to diversifying gifted education programs. Until that goal is reached, educators "should and can exercise greater vigilance in pursuit of educational values that will move alternative assessment reform in educationally rigorous, equitable, and sustainable directions" (Hargreaves, Earl, & Schmidt, 2002, p. 93).

References

Baker, E. L. (1996). Introduction to theme issue in educational assessment. *Journal of Educational Research, 89*, 194–196.

Bernal, E. M. (2002). Three ways to achieve a more equitable representation of culturally and linguistically different students in GT programs. *Roeper Review, 24*, 82–88.

Borland, J. M., & Wright, L. (1994). Identifying young, potentially gifted, economically disadvantaged students. *Gifted Child Quarterly, 38*, 164–171.

Clasen, D. R., Middleton, J. A, & Connell, T. J. (1994). Assessing artistic and problem-solving performance in minority and nonminority students using a nontraditional multidimensional approach. *Gifted Child Quarterly, 38*, 27–37.

Coleman, M. R. (2001). Middle schools: New trends and issues. *Gifted Child Today, 24*(4), 20–22.

Cummings, J. (1991). Institutionalized racism and the assessment of minority children: A comparison of policies and programs in the United States and Canada. In R. J. Samuda, S. L. Kong, J. Cummins, J. Pascual-Leone, & J. Lewis (Eds.), *Assessment and placement of minority students* (pp. 97–107). Toronto: Intercultural Social Sciences Publications.

Dunbar, S. B., Koretz, D. M., & Hoover, H. D. (1991). Quality control in the development and use of performance assessments. *Applied Measurement in Education, 4*, 289–303.

Eccles, J. S. (1993, March). *Psychological and social barriers to women's participation in mathematics and science.* Paper presented at the biennial meeting of the American Educational Research Association, Atlanta, GA.

Eccles, J. S., Wigfield, A., Midgley, C., Reuman, D., MacIver, D., & Feldlaufer, H. (1992). Negative effects of traditional middle schools on students' motivation. *The Elementary School Journal, 93*, 553–574.

Ford, D. Y., & Harmon, D. A. (2001). Equity and excellence: Providing access to gifted education for culturally diverse students. *Journal of Secondary Gifted Education, 12*, 141–148.

Ford, D. Y., Harris, J., Tyson, C. A., & Trotman, F. M. (2002). Beyond deficit thinking: Providing access for gifted African American students. *Roeper Review, 24*, 52–58.

Frechtling, J. A. (1991). Performance assessment: Moonstruck or the real thing? *Educational Measurement: Issues and Practices, 10*(4), 23–25.

Gabelko, N. H., & Sosniak, L. A. (2002). "Someone just like me": When academic engagement trumps race, class, and gender. *Phi Delta Kappan, 83*, 400–407.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: BasicBooks.

Griffiths, S. E. (1996). *The inter-observer reliability of the DISCOVER problem-solving assessment*. Unpublished manuscript, The University of Arizona, Tucson.

Hafenstein, N. L., & Tucker, B. (1994). *Performance-based assessment: An alternative assessment process for young gifted children*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans. (ERIC Document Reproduction Serves No. 370 994)

Hargreaves, A., Earl, L., & Schmidt, M. (2002). Perspectives on alternative assessment reform. *American Educational Research Journal, 39*, 69–95.

Lubinski, D., & Benbow, C. P. (1992). Gender differences in abilities and preferences among the gifted: Implications for the math-science pipeline. *Current Directions in Psychological Science, 1*, 61–66.

Lynch, M. E. (1991). Gender intensification. In R. M. Lerner, A. C. Patterson, & J. Brooks-Gunn (Eds.), *Encyclopedia of adolescence* (Vol. 1, pp. 395–397). New York: Garland.

Maker, C. J. (1992). Intelligence and creativity in multiple intelligences: Identification and development. *Educating Able Learners, 17*(4), 12–19.

Maker, C. J. (1993). Creativity, intelligence, and problem solving: A definition and design for cross-cultural research and measurement related to giftedness. *Gifted Education International, ,9*, 68–77.

Maker, C. J. (1996). Identification of gifted minority students: A national problem, needed changes, and a promising solution. *Gifted Child Quarterly, 40*, 41–50.

Maker, C. J., Nielson, A. B., & Rogers, J. A. (1994). Giftedness, diversity, and problem-solving. *Teaching Exceptional Children, 27*(1), 4–19.

Plucker, J. A., Callahan, C. M., & Tomchin, E. M. (1996). Wherefore art thou, multiple intelligences? Alternative assessments for identifying talent in ethnically diverse and low-income students. *Gifted Child Quarterly, 40,* 81–90.

Reid, C., Udall, A., Romanoff, B., & Algozzine, B. (1999). Comparison of traditional and problem solving assessment criteria. *Gifted Child Quarterly, 43*, 252–264.

Sarouphim, K. M. (1999). DISCOVER: A promising alternative assessment for the identification of gifted minorities. *Gifted Child Quarterly, 43*, 244–251.

Sarouphim, K. M. (2000). Internal structure of DISCOVER: A performance-based assessment. *Journal for the Education of the Gifted, 23*, 314–327.

Sarouphim, K. M. (2001). DISCOVER: Concurrent validity, gender differences, and identification of minority students. *Gifted Child Quarterly, 45*, 130–138.

Sarouphim, K. M. (2002) DISCOVER in high school: Identifying gifted Hispanic and Native American students. *Journal of Secondary Gifted Education, 14*, 30–38.

Sternberg, R. J., & Davidson, J. (1986). *Conceptions of giftedness*. Cambridge: Cambridge University Press.