The Use of a Performance Assessment for Identifying Gifted Lebanese Students: Is DISCOVER Effective?

Ketty M. Sarouphim
Lebanese American University

The purpose of this study was to investigate the effectiveness of DISCOVER, a performance-based assessment in identifying gifted Lebanese students. The sample consisted of 248 students (121 boys, 127 girls) from Grades 3–5 at two private schools in Beirut, Lebanon. Students were administered DISCOVER and the Raven Standard Progressive Matrices (RSPM). Teachers and parents of identified students were interviewed. Results showed evidence for DISCOVER's concurrent validity with RSPM, as correlations between students' DISCOVER ratings in spatial intelligence and their Raven scores were high, whereas correlations between students' DISCOVER ratings in linguistic intelligences and their Raven scores were low. Parents' and teachers' interviews corroborated the results, with a few exceptions. Of the total sample, 14.5% were identified, with no gender differences. The researcher concluded that DISCOVER could be used effectively in Lebanon for identification purposes, but further research is needed to support the findings.

The identification of gifted students has traditionally relied on the use of IQ testing. Scores in the 97th percentile or above have typically enabled students' entry into programs for the gifted (Ford, Harris, Tyson, & Trotman, 2002). However, in the last two decades, scholars and researchers have questioned the use of IQ tests for identification purposes, due to the mounting evidence that these measures do not constitute the sole indicators of giftedness (Baldwin, 2005; Maker, 1993); instead, alternative methods such as the use of performance-based assessment and portfolios were examined as possible additional or even replacement tools for identification purposes (Gallagher, 2005). One reason for the dissatisfaction with IQ tests is that their

use has led to the underrepresentation of students from economically disadvantaged and culturally diverse groups (Callahan, 2005), possibly because of their heavy reliance on verbal skills, which minority students often lack.

In Lebanon, a small country in the Middle East, a growing interest in the education of the gifted has been noted among educators and scholars even though at present, the country lacks a formal system for educating gifted students. For this reason, identification procedures constitute a major issue. In order to initiate programs in education of the gifted built on a solid basis, research on identification measures is much needed in the country, given that the growing interest in this field will eventually lead to the establishment of programs for the gifted in Lebanese schools. This study examines the effectiveness of DISCOVER (Maker, Nielson, & Rogers, 1994), a performance-based assessment in identifying gifted Lebanese students. DISCOVER is an acronym that stands for *D*iscovering *I*ntellectual *S*trength and *C*apabilities while allowing for *V*aried *E*thnic *R*esponses.

DISCOVER was chosen for this study because of its many characteristics that make it suitable for use with non-English-speaking populations. Although other tests could be useful with non-English-speaking students, DISCOVER has many advantages over other instruments, such as its reliance on manipulatives and the possibility of its administration in the native language of the children tested. Other advantages include the assessment of multiple intelligences and the enjoyment and engagement of children in taking the test.

Status of Education of the Gifted in Lebanon

Lebanon is one of the smallest countries in the Middle East (10,452 square kilometers), almost the size of the state of Connecticut. Schooling follows a lock-step system, with little attention given to individual differences. Programs for students with special needs are sparse, but when found their focus is mostly on students with academic problems rather than on students who show high ability. The closest to a formal program for the gifted is found in some private schools located in Beirut, the capital. These schools cater to students from high socioeconomic status and have established for high-achieving

students what is similar to enrichment programs on pull-out bases. These programs are, however, rudimentary and effected haphazardly without grounding in a solid theoretical base. To a large extent, the reason for the lack of programs for the gifted in the country is due to the lack of reliable and valid instruments for identification. Currently, the country lacks measures standardized on samples of Lebanese students to assess intelligence; the only tests used for this purpose are imported from the West and translated into Arabic, the native language of the Lebanese, a process dangerously flawed. Hence, a great need exists for reliable and valid instruments for the identification of gifted Lebanese students as well as for programs for the gifted built on grounded theory.

Giftedness can be found in all cultures and is expressed through a variety of behaviors (Baldwin, 2005). Parents, teachers, and school officials need to have knowledge in giftedness and its determinants so that they become "talent spotters," always on the lookout for untapped hidden ability. The first step in this process is to adopt a broad view of the concepts of intelligence and giftedness that goes beyond high academic performance and encompasses a wide range of human abilities. This is of particular importance in Lebanon where work on identification and programs for gifted students have just started, hence the significance of all kinds of research in this budding area and of this study in particular. One advantage of the lack of programs for gifted students in the country is the "clean slate" phenomenon; that is, educators can start working afresh, molding the field of education of the gifted based on empirical evidence yielded by the pool of research findings already available in the Western literature, a process potentially less problematic than attempting to fix flaws in preexisting programs. However, a thorough examination of the Western research is needed for fine-tuning of the findings to the Lebanese population instead of just importing evidence from the West and applying it blindly in the Lebanese culture. In sum, establishing a discipline of gifted education in Lebanon, effective and unique to the country, is timely and critical. In a country with a history troubled with repetitive wars, it is the civic responsibility of scholars and educators to recognize and nurture the talents of its gifted citizens who represent the untapped promise for a better future for Lebanon.

Research on DISCOVER

The use of performance-based assessments has been on the rise in the last two decades. Advocates cite many advantages for using these instruments, such as assessment of higher order skills, reducing the gap between testing and instruction, their coverage of broad areas of intelligence, and assessing students in life-like and complex situations (Maker, 1993; O'Neil, 1992). Several studies have been conducted on the effectiveness of these relatively new instruments, with mostly positive results (e.g., Pierce et al., 2007; Reid, Udall, Romanoff, & Algozzine, 1999; VanTassel-Baska, Johnson, & Avery, 2002), suggesting a paradigm shift in identification procedures. Madaus and O'Dwyer (1999) called performance assessments the "new technology" (p. 688) of testing today.

The performance-based assessment DISCOVER is grounded in Gardner's (1983) multiple intelligences (MI) theory and based on Maker's (1993) definition of giftedness: "the ability to solve the most complex problems in the most efficient, effective, or economical ways" (p. 71). Based on tasks (manipulatives) that require problem-solving and creative abilities, the instrument assesses six of the intelligences identified by Gardner—linguistic, spatial, logical-mathematical, interpersonal, intrapersonal, and bodily-kinesthetic—through activities carefully designed to tap into students' strengths and capabilities in these intelligences.

DISCOVER was developed to identify gifted students from culturally diverse groups and was tested in many countries where English is not the native language (Baldwin, 2005). Since its inception at the University of Arizona (Maker et al.,1994), DISCOVER has been administered to students in the United States and to thousands of students from diverse populations in countries such as China, Taiwan, France, Australia, England, and Bahrain (Maker et al., 2006). However, most research on the reliability and validity of the instrument has included data collected in the United States, as shown in the following reviews.

Interrater Reliability

In a triangulated inquiry on the interrater reliability of DISCOVER, Sarouphim (1999) investigated the alignment of ratings given to students by three independent raters: DISCOVER observers, the classroom teacher, and the researcher. The results showed that the DISCOVER observers, classroom teacher, and researcher gave similar ratings to students in the linguistic, spatial, and mathematical intelligences assessed in DISCOVER through structured activities, but their ratings were not as similar in the personal and bodily-kinesthetic intelligences assessed in DISCOVER through unstructured tasks. The researcher concluded that the DISCOVER observers were more effective in appraising students' intelligences when the appraisal was made through specific activities than when the appraisal depended on observing unstructured behavior. Sarouphim (1999) recommended that specific activities be developed for accurate appraisal through DISCOVER of the whole spectrum of multiple intelligences.

In another study on DISCOVER's interrater reliability, Giffiths (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 MI theory through a series of interobserver correlations. The sample consisted of 254 elementary students, predominantly from economically disadvantaged Native American and Hispanic groups. The results showed low interobserver correlations across grade levels between the activities that measure different intelligences (e.g., linguistic and spatial activities) and moderate to high correlations between activities that measure related intelligences (e.g., oral and written linguistic), indicating that students identified in one intelligence were not necessarily found gifted in the other intelligences. The results suggested that the different DISCOVER activities possibly

measure different intelligences, a finding that supports the consistency between DISCOVER and Gardner's MI theory.

Comparative Validity

Griffiths (1996) examined the comparative validity of DISCOVER with the WISC-III. The sample consisted of 30 Mexican American low-income children whose ages ranged between 9–11 years. Griffiths examined the relationship between students' ratings on each of the DISCOVER activities and their scores on the corresponding WISC-III subtests. Although overall students' ratings in the two measures were different (i.e., students identified as gifted through DISCOVER did not necessarily have IQ scores in the top 3%), analyses of separate activities corresponding to the different intelligences (e.g., math, linguistic, etc.) showed close resemblance, indicating evidence for the concurrent validity of DISCOVER with WISC-III.

Gender and Ethnic Differences

Finally, Sarouphim (2005) examined the effectiveness of DISCOVER on a large sample of 955 students taken from Grades K–12 in 10 schools in Arizona. The results revealed a good fit between DISCOVER and MI theory; also, no significant ethnic or gender differences in identification were found. A total of 20.9% participants were identified, suggesting that DISCOVER might contribute to diminishing the problem of minority underrepresentation in programs for the gifted.

In sum, research on DISCOVER has yielded mostly positive results on its effectiveness in identifying students from culturally diverse groups. As Baldwin (2005) stated: "[DISCOVER] has undergone revisions to make it a valid and reliable tool for identifying individual strengths in multiple intelligences" (p. 107). However, data used in research on DISCOVER were mostly collected in the United States. The only study that was conducted on DISCOVER in Lebanon, the country where the present investigation has taken place, has included a small sample and aimed at testing the ground for whether the assessment could be used effectively in Lebanon (Sarouphim, 2007). In that study, DISCOVER was administered to 49 fifth graders taken from one private school in Beirut. The results

were similar to those found in the Western populations and showed a similar pattern of identification trends, such as in the high percentage of students identified (19% of the sample) and the lack of gender differences in identification. The results were also corroborated by interviews with teachers and the students' grade reports, indicating that DISCOVER could be used effectively in Lebanon. In the current study, a larger sample was used as well as additional instruments to validate the data. Three research problems were investigated in this study:

- The concurrent validity of DISCOVER with the Raven's Standard Progressive Matrices (RSPM).
- Congruence of the DISCOVER ratings with those of parents and teachers, as well as with the students' school grades.
- Gender differences in identification.

Method

Participants

The sample consisted of 248 students (121 boys and 127 girls) taken from Grades 3–5 at two private schools in Beirut, Lebanon. The participants were predominantly from the middle SES, as evidenced by their place of residence and their parents' occupation. The participants' mean age was 8.1 years in third grade ($n = 86$, 36 boys and 50 girls), 9.2 years in fourth grade ($n = 77$, 43 boys and 34 girls), and 10.2 years in fifth grade ($n = 85$, 42 boys and 43 girls). They were all Caucasians and spoke English as well as Arabic.

Procedures

All participants were given the Grades 3–5 version of the DISCOVER assessment followed by the Raven Standard Progressive Matrices (RSPM). The administration of both tests took place in the students' classrooms on two consecutive days. Both the researcher and the observers were well-trained in the use of both instruments. Interviews with the homeroom teachers and with the parents who nominated

their children were conducted following the administration of the two instruments. Interviews took place after school hours and were tape recorded with the permission of the interviewees. The criteria for identification were set according to the standards specified by the developers of DISCOVER: a rating of *Definitely* in at least two of the assessment's activities.

Instruments

The Raven Standard Progressive Matrices. The RSPM is a test of nonverbal ability designed to be equally effective with English and non-English-speaking populations. The RSPM is now used internationally for comparative purposes (Raven, Raven, & Court, 1998), thus the choice for its use in this study. Another reason for its use in this current investigation is that the RSPM is considered to be relatively culturally fair (Chaffey, Halliwell, & McCluskey, 2006) and has been tested with different populations. Basically, the test comprises 60 problems (divided into 5 sets of 12 items each) and consists of presenting a series of figures, each with a missing piece. Below each figure are six to eight alternative pieces, but only one completes the figure. The figures are presented in ascending difficulty and keep the student motivated (Baldwin, 2005; Raven et al., 1998).

Both the reliability and validity of RSPM are reasonably high. Split-half internal consistency coefficients reported in the literature exceed .90, with a modal value of .91 (Raven et al., 1998). Test-retest reliability coefficients range between .78 to 1.00, depending on the interval between tests, with higher reliability correlations associated with shorter intervals. In terms of validity, studies on RSPM with concurrent intelligence measures, such as Binet and Wechsler Scales, have yielded a coefficient range from .54 to .86 (Raven et al., 1998). However, the major limitation of the instrument is that it measures one kind of ability, namely nonverbal reasoning, to the exclusion of the other intelligences.

The DISCOVER Assessment. DISCOVER is performance-based and includes tasks that increase progressively in complexity and openness. Like the RSPM, the assessment is not a traditional test and children have fun working on solving the assigned problems. Basically, three activities are performed in class during the

DISCOVER administration to assess spatial, mathematical, and oral linguistic intelligences. Logical-mathematical and written linguistic intelligences are measured a day or so following the classroom assessment through paper-and-pencil tasks. Bodily-kinesthetic and the personal intelligences are assessed by observing the behaviors of students throughout the administration, which typically lasts 2 ½ hours.

The DISCOVER assessment measures the different intelligences by using separate activities across intelligences and age levels. Different tasks are designed for grade levels from Kindergarten through Grade 12. The focus in this study is on the DISCOVER version for Grades 3–5. The following is a brief description of the activities in this version. (For a detailed account, see Maker, 1992).

Spatial artistic. Students are provided with colored cardboard pieces of different shapes, designs, and sizes and asked to make different constructions with these pieces (e.g., geometrical designs, flowers, something that moves, and a free construction).

Spatial analytical. 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 first requested to make a triangle 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" that consist of more difficult puzzle problems.

Oral linguistic. Students are provided with an array of toys and asked to provide as many descriptors of the items as possible. Then each student is asked to tell a story of his or her choice that incorporates some or all of the provided toys.

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

Logical-mathematical. 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 "problem-finding," that is, creating as many problems as possible (addition, subtraction, etc.) with a prespecified number as the answer.

Interpersonal, intrapersonal, and bodily-kinesthetic. Although these intelligences are not measured through specific

activities, behaviors corresponding to students' strengths in these intelligences are noted by the observers. For example, statements such as "I can't give up now, I know I can solve this puzzle" is considered strength in intrapersonal intelligence; cooperative behavior in the form of helping a classmate to finish a task is considered strength in the interpersonal intelligence, and incorporating one's own body into a construction or forms of graceful movements are noted as strengths in bodily-kinesthetic intelligence.

Typically, the DISCOVER administration takes place in the classroom. Trained observers gather around the children with an approximate ratio of 1:5 (one observer to five children). Each observer takes notes and records observed behaviors on standard sheets while the classroom teacher gives instructions in the children's dominant language. Observers pay attention to the children's problem-solving process as well as to their products. To avoid observer bias, observers rotate at the completion of each activity; thus, each child is observed by at least two persons during the administration.

Following the administration, all observers meet to discuss the students' strengths and complete a behavior checklist on each child. Observers classify children's strengths in each activity into one of four different categories ranging from "no strength observed" to a "definite strength observed" using the ratings of *Unknown*, *Maybe, Probably,* and *Definitely*. The category *Definitely* corresponds to high ability or to giftedness in that particular intelligence assessed by its corresponding activity. A child who gets a *Definitely* rating in at least two of the activities is identified as gifted.

As described in the review of literature, studies on the properties of DISCOVER have yielded mostly positive results. However, a major limitation of the instrument is that three of the intelligences—interpersonal, intrapersonal, and bodily-kinesthetic—are not assessed through specific activities, which might affect the objectivity of the ratings. Another limitation that might impact this study is that the tasks included in the DISCOVER activities might not be familiar to Lebanese students, thus possibly impacting their performance and consequently increasing the margin of error in their ratings.

Data Analysis

Separate but identical analyses were performed on the data collected in each grade; then the same analyses were performed on data pooled from the entire sample. To answer the first research question on concurrent validity, the students' DISCOVER ratings were correlated with their RSPM scores, thus yielding correlation coefficients on the concurrent validity of RSPM with each of the DISCOVER activities.

For the second research question on the congruence of the students' DISCOVER ratings with their school performance, interviews were conducted with the teachers and students' parents. All interviews were transcribed and coded. Also, the students' DISCOVER ratings were matched with their grades in the subject areas corresponding to the DISCOVER tasks, namely math, art, and language.

Finally, to answer the third research question on gender differences, a 2 x 3 MANOVA was performed (Gender X Grade). MANOVA was used because the two variables of gender and grades (independent variables) were investigated with regard to the students' performance in the five DISCOVER activities (the dependent variables). The aim was to examine whether significant differences existed between the ratings given to boys and those given to girls in the activities measuring each of the intelligences. The ratings were coded as follows: 1 for *Unknown,* 2 for *Maybe,* 3 for *Probably,* and 4 for *Definitely.* According to the DISCOVER criteria for identification, students given the rating of *Definitely* in two of the activities are considered gifted. Therefore, the number of boys and girls who fit this criterion was calculated, then chi-square tests of significance were performed for gender by gifted participants to examine gender differences in identification.

Results

Concurrent Validity

In this study, the students' ratings in each of the DISCOVER activities were correlated with their RSPM scores. As shown in Table 1, the highest correlation was found between the students' ratings in

Table 1

Correlations Between Raven Scores and DISCOVER Ratings

	Third Grade ($n = 86$)		Fourth Grade ($n = 77$)		Fifth Grade ($n = 85$)		All ($N = 248$)	
	r	p	r	p	r	p	r	p
Spatial Artistic	.418**	0.01	.551**	0.00	.603**	0.00	.547**	0.01
Spatial Analytic	.613**	0.00	.777**	0.00	.824**	0.00	.776**	0.00
Math	.269*	0.05	.311*	0.04	.427**	0.01	.454**	0.01
Oral	.212	0.10	.198	0.19	.224	0.09	.201	0.11
Written	.127	0.22	.072	0.54	.137	0.20	.114	0.23

*Note. *$p < .05$. **$p < .01$.*

the spatial analytic activity and their RSPM scores ($r = .776, p < .01$) followed by that between their ratings in the spatial artistic activity and their RSPM scores ($r = .547, p < .01$). The lowest correlations were between the students' RSPM scores and their ratings in the linguistic activities, oral linguistic ($r = 0.2, p = 0.11$, ns), and written linguistic ($r = .114, p = 0.23$, ns).

Parents' and Teachers' Nomination

The results showed that students identified through the use of DISCOVER were all nominated by either their parents, teachers, or both. As one teacher said: "I knew all along that [these students] were gifted. I am glad to know that my hunch was validated by a formal assessment." However, 7 students nominated by parents and 3 by their teachers were not identified through the use of DISCOVER. Interviews with the parents who disagreed with their children's scores showed that they believed strongly in their children's high ability, especially spatial artistic ability, a belief that was not substantiated by their DISCOVER ratings. As one parent stated: "When I was told that my child was being given nontraditional assessments, I thought that he finally had a chance to shine; I am disappointed because I don't think that these tests did justice to my son's true capability." Another mother said

> I always thought that my daughter was a born artist. Her artistic skills started to show as early as when she was 3 years old. I remember that she drew then a house with a garden in perfect proportions and with all the elements present. She spends most of her free time drawing. When people see her drawings, they cannot believe that this is the work of a child. I don't understand how these alternative assessments did not reveal her high artistic ability.

The teachers' interviews mostly corroborated the results; that is, teachers mostly agreed with DISCOVER's findings in relation to the identified participants, except in the case of 3 students, all males in fourth grade, and all in one section. The teacher expressed her disappointment with the results and stated that all 3 nonidentified students deserved to be given the label "gifted." When asked to justify her comment, she said:

> I have watched these students struggle with school work since the first day of class, but somehow I thought that the three of them were highly gifted in many different ways even though their grades were never outstanding. It's just a feeling that I had, based on their good behavior and great effort. I still believe that their abilities will be "discovered" one day, even if DISCOVER did not discover that!

Alignment of Students' Grades With Their DISCOVER Ratings

All identified students had high grade point averages, ranging between 3.8 and 4.0, indicating that their DISCOVER ratings were aligned with their classroom performance. Interviews with the teachers revealed that the identified students had been getting high grades throughout the school year and accordingly, "they deserved to be given the label gifted."

Table 2

*Mean Ratings of Students in the
DISCOVER Activities Across Grades*

		Third Grade		Fourth Grade		Fifth Grade	
		M	F	M	F	M	F
Spatial Artistic							
	M	3.1	2.9	3.0	2.8	2.9	2.8
	SD	1.1	0.9	0.8	0.6	0.5	0.4
Spatial Analytic							
	M	3.0	3.0	2.6	2.7	2.8	2.7
	SD	0.2	0.7	1.0	0.9	0.5	0.8
Logical-Math							
	M	2.7	2.5	2.6	2.5	3.1*	2.4
	SD	0.7	0.5	0.8	0.9	0.4	0.3
Oral Linguistic							
	M	2.9	3.1	3.0	3.0	2.8	2.9
	SD	0.7	0.9	1.2	0.8	1.0	0.8
Written Linguistic							
	M	2.8	2.7	2.9	2.8	2.8	2.6
	SD	1.0	0.5	0.6	0.7	0.4	0.4

Note. Unknown = 1; Maybe = 2; Probably = 3; Definitely = 4.
*$p < .05$.

Gender Differences in Identification

A total of 36 children or 14.5% of the sample (19 boys and 17 girls) were identified through DISCOVER. Although the boys' DISCOVER ratings were generally slightly higher than those of the girls (see Table 2), the results showed no significant gender differences in students' ratings across grade levels and activities, with one exception found in the math ratings of fifth graders. The 2 x 3 MANOVA yielded nonsignificant results for gender by grade interaction ($F[4,232]$) = 1.62, $p = 0.57$, ns), and no significant main effect for grade was found ($F[3,208] = 0.83$, $p = 0.67$, ns). However, a significant main effect for gender ($F[4,232] = 4.89$, $p = 0.03$) was found, eta-squared = 0.36,

Table 3

Chi-Square Tests of Significance for Gender by Gifted Participants

	Gifted Boys		Gifted Girls		Total				
Grade	*n*	%	*n*	%	*n*	%	*df*	χ^2	*p*
Third	8	16.6	6	16.0	14	16.3	1	0.09	0.34
Fourth	5	13.9	6	14.7	11	14.2	1	1.23	0.28
Fifth	6	12.0	5	13.9	14	12.9	1	2.21	0.09
Total	19	14.0	17	14.9	36	14.5	1	1.59	0.11

which led to further univariate analysis. The ANOVA revealed that boys significantly outperformed girls in math in fifth grade (F[1, 122] = 6.54, p = 0.01), with a moderate effect size of 0.41. Moreover, as shown in Table 3, chi-square tests did not reveal any significant gender differences in identification (i.e., no significant differences in the numbers of boys and girls who were given the *Definitely* rating in at least two of the DISCOVER activities) in all grade levels and across the entire sample (χ^2 = 1.59, p = 0.11, ns).

Discussion

This study examined the effectiveness of DISCOVER, an alternative assessment, in identifying gifted Lebanese students. Concurrent validity of DISCOVER with the RSPM was studied, as well as gender differences in identification. Also examined was the congruence between students' DISCOVER ratings and those of their parents and teachers on the one hand, and between students' DISCOVER ratings and their academic performance on the other. The results provided mostly positive results, in support of the use of DISCOVER in Lebanon.

The correlations between students' DISCOVER ratings and their RSPM scores showed that the spatial activities of DISCOVER were aligned with RSPM, indicating that the two measures assess similar abilities. Along the same lines, the results showed low correlations between students' RSPM scores and their ratings in the linguistic

activities of DISCOVER, suggesting that the two measures assess different abilities. These findings provide evidence for the convergent and discriminant validity of DISCOVER. Given that the RSPM is a measure of nonverbal ability, the high correlations found between students' RSPM scores and their DISCOVER ratings in the spatial activities suggest a high concurrent validity between the RSPM and DISCOVER. Similarly, the low correlations found between students' RSPM scores and their DISCOVER ratings in the linguistic activities give evidence for discriminant validity. These results also suggest that a good fit exists between the assessment's internal structure and the theory of multiple intelligences because students who were given a high rating in the spatial activities were not necessarily given high ratings in the linguistic activities as well, possibly indicating that the spatial and linguistic activities measure separate and different intelligences. Similar results were obtained in research conducted on samples of American students (Sarouphim, 2001, 2002, 2005).

Another interesting finding is the significant gender differences found in students' math ratings in fifth grade, but not in the younger participants. In current Western research, the results show that the gender gap in math has been narrowing, although previous findings indicated that gender differences in math start showing in late elementary school, possibly because of the "masculine" label associated with high mathematical ability (Hyde, 2005). Early adolescence is the age during which individuals try to define their values and gender roles. Because the culture in Lebanon tends to be fairly traditional, it is not surprising that the boys' ratings in math were found to be superior to those of the girls, as Lebanese early adolescent girls do not wish to risk being labeled masculine. Similar results were found in the research performed on DISCOVER encompassing samples of American students from culturally diverse groups (Sarouphim, 2001).

Along the same lines, another noteworthy finding is the lack of significant gender differences in identification, suggesting that the use of DISCOVER does not produce gender bias. Similar results were obtained in the pilot study conducted in Lebanon (Sarouphim, 2007). Educators have long shunned the use of instruments that lead to any kind of discrimination, especially with regard to gender or race. The fact that DISCOVER did not lead to gender identification bias in a culture as traditional as the Lebanese culture is quite significant.

Programs for the gifted in Lebanon should be established on solid grounds; hence, it is mandatory that the instruments for identification used be free of gender bias from the beginning.

An interesting but not surprising result is that of the discrepancy found between a few of the teachers' and parents' nominations on the one hand and the students' corresponding DISCOVER ratings on the other. This finding is not surprising because several studies have documented the misguided perceptions of parents and teachers on what they consider evidence for giftedness (Baldwin, 2005; Pierce et al., 2007). More often than not, teachers develop their own theories of what is gifted behavior and at times confuse good behavior with giftedness. Similarly, parents tend to overestimate their children's abilities, particularly their sons' (Miller, 2006). On the other hand, the result that all students identified as gifted were nominated by their teachers is in fact unexpected and does not corroborate findings in the literature (Baldwin, 2005). Typically, teachers underestimate the abilities of students who do not fare well academically. Moreover, teachers consider themselves more knowledgeable in their students' abilities than any instrument could reveal. As one teacher explained, "Teachers work with the child on daily basis. They can claim to know the child better than any test does." One explanation for the good match found in this study between teachers' nominations and the DISCOVER ratings may be due to the fact that all identified students had high grade point averages, thus confirming the teachers' belief that giftedness is interchangeable with high grades. In Lebanon, teachers do not have much experience, if any, in educating the gifted. Both teachers and parents need to be educated in what constitutes giftedness, given that education of the gifted is a new field in the country.

Another finding in support of the use of DISCOVER in Lebanon is that related to the alignment between the students' DISCOVER ratings and their school grades. Similar results were obtained in the pilot study conducted on a small sample of Lebanese students (Sarouphim, 2007). Although this finding is unexpected and does not necessarily corroborate the results of previous research conducted in the West, which have mostly showed that students identified through DISCOVER do not all score high on traditional standardized tests (Sarouphim, 2005), the finding is nevertheless significant, because no Lebanese instrument currently exists for the identification of gifted

students. One would want to start a program for the gifted with an instrument that provides a good fit with school grades; otherwise, a clash between the two might affect negatively the validity of the instrument. In this regard, one explanation for the difference in results found in this study and those in studies conducted in the West might be related to the better match between the abilities needed to solve the DISCOVER tasks and those in the curriculum adopted in Lebanese schools. For example, a great emphasis is placed in Lebanese schools on the mastery of foreign languages, mostly English and French (Ghaith, 2003). By the time they graduate from high school, Lebanese students are almost as fluent (and in some cases even more fluent) in at least one foreign language as they are in their own native Arabic language. This emphasis might explain the high ratings given to study participants in the linguistic activities. Another explanation is that the Lebanese culture is a collectivistic culture as opposed to the individualistic culture of the West. It is a tight-knit culture in which traditions are transmitted orally from one generation to the next in the form of stories and folk tales. Thus, storytelling (a task required in the oral linguistic activity of DISCOVER) is integral to the Lebanese culture and is engrained in its daily life, which might explain why Lebanese children fared better in the linguistic activities than the Western sample did.

A finding that also does not corroborate the results of Western research is the lower percentages of students identified in Lebanon, about 25% less than those identified in the American sample (Sarouphim, 2001, 2005). Although Lebanese students did better in the linguistic activities of DISCOVER, they did not do as well in the math and analytical activities, possibly because the tasks required in these two activities are less related to the material taught in Lebanese schools. For example, one of the problems in the DISCOVER math worksheet involves solving magic squares, with which American students are quite familiar, but for many Lebanese students, coming across this task in DISCOVER might be their first encounter with solving magic squares.

Limitations and Recommendations for Future Research

One major limitation of this study is that the sample used was restricted to two private schools in Beirut, thus the results cannot be generalized to the Lebanese student population. Another related limitation is that the participants were all third, fourth, and fifth graders, which confines the findings to elementary school students. Therefore, in future research, larger samples encompassing a wide range of students taken from different geographical locations in Lebanon must be included.

Furthermore, a third limitation is that DISCOVER was originally designed for use with students in the United States, therefore the validity and reliability of this instrument must be established with samples of Lebanese students before one could sound a call for its use on a large scale in Lebanon. Consequently, in future research, emphasis must be placed on examining the DISCOVER tasks with scrutiny to determine the necessary modifications to be made to the Lebanese version without compromising the integrity of the assessment or its original design.

In conclusion, DISCOVER might become a leading instrument used for the identification of gifted students in Lebanon, but further research is of utmost importance, especially because education of gifted students is a relatively new area in the country and needs to be established on a solid basis.

References

Baldwin, A. Y. (2005). Identification concerns for gifted students of diverse populations. *Theory Into Practice, 44,* 105–114.

Callahan, C. M. (2005). Identifying gifted students from underrepresented populations. *Theory Into Practice, 44,* 98–104.

Chaffey, G. W., Halliwell, G., & McCluskey, K. W. (2006). Identifying high academic potential in Canadian Aboriginal primary school children. *Gifted and Talented International, 21*(2), 61–70.

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

Gallagher, J. J. (2005). The role of race in gifted education. *Roeper Review, 27,* 135–136.

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

Ghaith, G. (2003). The relationship between forms of instruction, achievement and perceptions of classroom climate. *Educational Research, 45,* 83–93.

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

Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist, 60,* 21–31.

Madaus, G., & O'Dwyer, L. (1999). Short history of performance assessment: Lessons learned. *Phi Delta Kappan, 9,* 688–695.

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., Muammar, O., Serino, L., Kuang, C. C., Mohamed, A., & Sak, U. (2006). The DISCOVER curriculum model: Nurturing and enhancing creativity in all children. *Journal of Educational Policy, 3,* 99–121.

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

Miller, E. M. (2006). Characteristic centrality in the perceptions of giftedness as a predictor of the pattern of nomination of students for placement in gifted programming. *Roeper Review, 28,* 179–181.

O'Neil, J. (1992). Putting performance assessment to the test. *Educational Leadership*, *49*(8), 14–19.

Pierce, R., Adams, C., Speirs Neumeister, K., Cassady, J., Dixon, F., & Cross, T. (2007). Development of an identification procedure for a large urban school corporation: Identifying culturally diverse and academically gifted elementary students. *Roeper Review, 29,* 113–118.

Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for the Standard Progressive Matrices*. London, England: H. K. Lewis.

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

Sarouphim, K. M. (2005). DISCOVER across the spectrum of grades: Identifying gifted minority students. *Gifted and Talented International, 20,* 70–77.

Sarouphim, K. M. (2007, April). *DISCOVER in Lebanon: A pilot study.* Paper presented at the Annual Meeting of the American Educational Research Association, Chicago. (ERIC Document Reproduction Services No. ED496155)

Sarouphim, K. M. (1999). DISCOVER: A promising alternative assessment for the identification of gifted minorities. *Gifted Child Quarterly, 43,* 244–252.

VanTassel-Baska, J., Johnson, D, & Avery, L. (2002). Using performance tasks in the identification of economically disadvantaged and minority gifted learners: Findings from project STAR. *Gifted Child Quarterly, 46,* 110–124.