

---

# Group-Specific Norms and Teacher-Rating Scales: Implications for Underrepresentation

Journal of Advanced Academics  
23(2) 125–144  
© The Author(s) 2012  
Reprints and permission:  
sagepub.com/journalsPermissions.nav  
DOI: 10.1177/1932202X12438717  
<http://joaa.sagepub.com>



Scott J. Peters<sup>1</sup> and Marcia Gentry<sup>2</sup>

## Abstract

For more than 30 years, underrepresentation of certain racial, cultural, and income groups in gifted and talented programs has been documented as a serious problem. Not only does this issue make gifted programs appear as if they are designed solely for upper class, dominant-culture individuals, but it also means that talented students from diverse backgrounds are not having their needs met. This study sought to determine the utility of applying group-specific norms to achievement tests in identifying more proportional numbers of gifted and talented students from low-income families. In addition, this study applied the use of a teacher-rating scale to locate even more underserved students with high potential. Results indicated that both practices helped identify more underserved students than did using traditional general norm-group comparisons. Implications for identification policy and practice are shared.

## Keywords

gifted, achievement, identification, measurement, underrepresentation

The underidentification and resulting underrepresentation of students from certain racial, ethnic, and cultural groups in gifted and talented programs have been well documented since the 1970s (Donovan & Cross, 2002; Yoon & Gentry, 2009). Unfortunately, despite some progress concerning representation of students from diverse

---

<sup>1</sup>University of Wisconsin, Whitewater, USA

<sup>2</sup>Purdue University, West Lafayette, IN, USA

## Corresponding Author:

Scott J. Peters, Department of Educational Foundations, University of Wisconsin–Whitewater,  
800 West Main Street, Whitewater, WI 53190, USA

Email: [peterss@uww.edu](mailto:peterss@uww.edu)

cultures, students from low-income families remain underrepresented in gifted and talented programs (Stambaugh, 2007; Swanson, 2006) due partly to the strong correlation between academic achievement and family income (Valencia & Suzuki, 2001). These students remain underidentified despite the fact that 3.4 million students in the United States achieve in the top quartile and also come from families earning less than the median income (Wyner, Bridgeland, & DiIulio, 2009). The same authors found that more than 1 million of these students come from families whose incomes are so low that they qualify for the federal free and reduced lunch program. Two suggestions to solving this problem include the use of teacher-rating scales/behavior checklists (Stambaugh, 2007; VanTassel-Baska, 2008) and the use of local and group-specific norms when using achievement or aptitude data for student identification (Lohman, 2006).

Exploring the usefulness of combining these two suggestions in helping to locate gifted and talented students from low-income families was the primary focus of this study. Using an individual school data set, we explored the utility of combining local group norms and teacher ratings as an illustrative case study. Although the use of local and group-specific norms can be applied to a variety of different types of assessment data and content areas, here we focused on reading as measured by one state's achievement test. We also used several different percentages and cut scores, which could be changed when implemented in a specific school context.

## **Literature Review**

### *High-Achieving Students From Low-Income Families*

The relationship between family income and academic achievement is strong and pervasive (Viadero, 2006). Of the small group of low-income students who are high achieving in first grade (7% of low-income students), only just over half (56%) will remain high achieving by fifth grade (Wyner et al., 2009). In addition, the same authors found that these students tend to drop out of high school at a rate twice as high as their high-achieving, higher income peers. With regard to closing achievement gaps, high-achieving students from low-income families continue to grow academically and learn at rates much slower than their higher income peers (Plucker, Burroughs, & Song, 2010). In their "Mind the Other Gap" report, Plucker et al. (2010) found that the percentage of students from low-income families scoring at advanced levels increased by 0.8% in Grade 4 reading and by 0.3% in Grade 8 reading compared with growth rates of 1.2% in Grade 4 reading and 3.3% in Grade 8 reading for their higher income peers. High-achieving students from low-income families increased in representation in the advanced group by 0.8% in Grade 4 reading and 0.3% in Grade 8 reading, whereas their higher income peers increased in representation by 1.2% and 0.6%, respectively. These rates occurred over approximately 10 years.

By considering these facts and figures, two points become clear. First, high-achieving students who also come from low-income households are not being identified and served in programs meant to foster, develop, and encourage high achievement (such as

gifted and talented programs) even when they are explicitly high achieving. In addition, the figures above likely did not include underachieving students from low-income families. Reis and McCoach (2000) indicated these underachievers comprise a large group of students who remain unidentified and unserved in gifted programs. Second, for those students who are identified using ability or achievement test scores, the identification measures typical of gifted and talented programs (National Association for Gifted Children [NAGC], 2009), their skills go underdeveloped as indicated by the decreasing rates of high achievement among these students beginning as early as elementary school. These students are progressing at much slower levels when compared with their peers from higher income families (Plucker et al., 2010).

### *Teacher-Rating Scales<sup>1</sup>*

A great deal of debate exists over the utility and appropriateness of using teacher-rating scales in the identification of gifted and talented students. Teacher ratings, referrals, and nominations have often been criticized for their lack of validity evidence (e.g., Pagnato & Birch, 1959; Peterson & Margolin, 1997). Pagnato and Birch's (1959) article is one of the most widely cited with regard to an empirical basis for such a conclusion. In their study, the authors used teacher nominations for giftedness in several content and nonacademic areas (e.g., art, music, social/political). The 154 students who were nominated as "mentally" gifted were then assessed using the Stanford-Binet Intelligence Test. Only 91 of the 154 students who were nominated obtained an IQ score of 136 or higher. This cut-point was used because it yielded the top 1% of the population on this IQ test. Based on this finding, the authors proposed that teacher nominations were neither efficacious nor efficient. However, when Gagné (1994) reanalyzed the same data, he found that with regard to student identification, "Teachers do not come out worse than most other sources of information" (p. 126), including mental ability tests, school grades, and achievement test scores; thereby, calling into question whether teachers really were as inaccurate as Pagnato and Birch had suggested.

Hodge and Cudmore (1986), in their extensive review of the literature concerning teacher-judgment measures, concluded that "the use of teacher judgments in the identification of gifted children should be continued, and, in fact, expanded" (p. 192). In reviewing 22 studies from 1959 through 1983, the authors recommended that to make teacher nominations a valid means of identification, the nominators need explicit definitions of giftedness in addition to well-developed instruments to help express their judgments and observations. Importantly, they concluded that just as much, if not more, research exists to support the use of teacher nomination scales as exists against their use.

Finally, with regard to the specific identification of low-income students, Stambaugh (2007) and VanTassel-Baska (2008) suggested that teacher-rating instruments have potential as screening tools or as an additional pathway to identification for students who are often overlooked. Peterson (1999) and High and Udall (1983) suggested that

teachers are in a prime position to recognize talent in their classrooms. From this perspective, researchers see teachers as evaluators of student performances over time and in classroom contexts as a valuable performance assessment for identifying students for gifted and talented programs.

The success of teachers as raters stands out thematically in the literature when teachers use specific criteria or instruments, and when they rate explicit behaviors. Although we found no studies comparing the efficacy of using a teacher-rating scale to a general teacher nomination to identify students, the emphasis on explicit behaviors as opposed to general teacher opinions and perceptions is recurrent in the literature. Furthermore, when teachers are simply asked for names of students whom they believe are gifted, the resulting nominations do not appear to be especially accurate. However, validity studies conducted using an array of published teacher-rating instruments indicate moderate to high criterion-related validity evidence. For example, concurrent validity coefficients for the *Gifted Rating Scales* (GRS; Pfeiffer & Jarosewich, 2003) are as high as .62 (GRS-S) when compared with an individually administered achievement test. Although the authors of the *Scales for Identifying Gifted Students* (SIGS; Ryser & McConnell, 2004) did not correlate scale scores with those obtained on achievement tests, they reported concurrent validity coefficients with three different ability/aptitude tests ranging from .38 to .89. These results suggest that teachers can be successful in identifying students, at least when correlated with outcomes such as achievement and ability tests. However, it is important that concurrent correlations between a teacher-rating scale and an achievement or aptitude test are moderate. If these correlations are too high, then the teacher-rating scale is simply duplicating the same information as the achievement test without contributing anything new.

### *Local and Group-Specific Norms*

One method that has been presented in the gifted and talented literature, but that has yet to see widespread application or implementation, is the use of local and/or group-specific norms. In the case of students from low-income families, the argument for using local norms is based on the idea that if schools are going to foster the talents and skills of all students, definitions of "high achieving" cannot be based on national norms (Lohman, 2005, 2006). When students from low-income families are represented in national normative samples, these students are unlikely to obtain the high scores required for gifted identification due to the strong, positive correlation between family income and academic achievement (Wyner et al., 2009) as well as inherent biases in many achievement and ability tests toward dominant middle-class students (Miller, 2004). In addition, the context and achievement make up of a given school could easily be very different from the national normative sample of any given test or instrument. Although this does not negate a test's quality, it does raise questions concerning which group should serve as the referent group. Lohman (2006) used the example of the 5% of American schools whose average achievement score is at the 95th percentile when

compared with national norms. In this case, identification decisions based on national norms would make little sense, as nearly every student would be identified as gifted. Although an appealing problem to have, identifying such a large number of students as gifted is unlikely to benefit them from the standpoint of programming or from a focus on student growth. Instead, programs and services need to be designed for those students whose needs are not met in the “general” classroom or by regular educational practices, whatever these look like in each context. In the case of the above-referenced school, a “gifted” program would look very different from the general educational experiences provided for the majority of the school’s students.

This same philosophy and practice of using local school norms can be applied to specific groups within those schools. If the desire is to find and nurture the gifts and talents of students from all socioeconomic strata, then group-specific norms should be used to locate top achievers within income or other subgroups of interest. This practice would identify those students from similar economic backgrounds who have achieved and who have needs not currently being addressed. For example, instead of identifying an arbitrary percentage such as the top 10% of all students in a given school as “the gifted,” administrators might instead select the top 10% of achievers within the group of students eligible for free or reduced lunch and from those students not eligible for free and reduced lunch. This practice, by controlling some degree of previous opportunity to learn, will better locate students who are high achieving but due to economic circumstances are not as high achieving as their higher income peers. This philosophy has been proposed in mainstream psychological testing books (Anastasi & Urbina, 1997) and is already used in at least one state, as described in recent studies by Matthews (2007) and Matthews and Kirsch (2011). In both studies, the authors discuss how Florida schools used different cut scores for some underrepresented groups than they did for dominant groups in gifted education identification. An important distinction is that the different cut scores used in Florida were not based on student performance (local norms) but were arbitrarily fixed beforehand. Such a practice would maintain the problem discussed earlier where entire schools or even no students in a given school might be identified as gifted. In a similar fashion to the Matthews (2007) and Matthews and Kirsch (2011) studies, a primary goal of the present study was to investigate whether or not local, group-specific norms, particularly related to family income, would help increase the representation of low-income students in gifted education programs. This practice is not without controversy as it involves differential cut scores based on group membership. However, this issue could be addressed by identifying all students who meet the general normative cut score (general norms), while also identifying all of the underrepresented students who meet a *group-specific* normative cut score (group-specific norms). This method would identify all those top-performing students, regardless of economic background, as well as those underrepresented students who show high achievement relative to their normative group. In this fashion, “gifted” becomes less of a categorical label, but instead focuses on developing the talents of all students, regardless of racial, ethnic, or socioeconomic background.

It is important to note that many of the most common definitions of giftedness and talent include specific language supporting the use of local and group-specific norms. For example, the *National Excellence Report* (United States Department of Education [U.S. DOE], 1993) presented the following definition:

Children and youth with outstanding talent perform or show the potential for performing at remarkably high levels of accomplishment *when compared with others of their age, experience, or environment*. These children and youth exhibit high performance capability in intellectual, creative, and/or artistic areas, possess an unusual leadership capacity, or excel in specific academic fields. They require services or activities not ordinarily provided by the schools. *Outstanding talents are present in children and youth from all cultural groups, across all economic strata*, and in all areas of human endeavor. (p. 3, italics added)

This issue is important for two reasons. First, because general education services and standards vary widely across all public schools, local norms should be used to determine which students need additional services beyond what is ordinarily provided. Second, the U.S. DOE definition (1993) cited earlier requires specific comparisons of students to others of similar background, experience, or environment. Making general, non-group-specific normative comparisons fails to meet this requirement. Instead, more targeted comparisons need to be made that take into account factors such as socioeconomic status or cultural group.

*Research Question 1:* How does the use of local, group-specific norms within a single school affect the representation of students from low-income families in gifted and talented programs? More specifically, how does use of these norms from an achievement test and from a teacher-rating scale work separately and together to identify students from low-income families?

## Method and Procedure

### Participants

Because of the focus on local and group-specific norms, a single, diverse school was purposively chosen to serve as an example and data source for this study. This school was not chosen to be widely representative or generalizable, but rather to serve as a case study in possible local-school identification practices. Teachers in this school had completed the HOPE Scale on their K-5 students as part of an earlier study (Peters & Gentry, 2010). The site, a K-8 public school located in a large, urban city in the Midwest, was specifically chosen not only for its overall diverse setting but also because its population of students eligible for free and reduced lunch (43%) matched that of the United States as of the 2007-2008 school year (42.9%; National Center

**Table 1.** Student Demographic Information (*N* = 539)

Designation	Percentage of sample
Free or reduced lunch program	43
Caucasian	33
African American	17
Hispanic	36
Asian	12
Multiracial	0
Native American	2
Gender (male)	48

for Educational Statistics [NCES], n.d.). Student demographics for this school are presented in Table 1. A final consideration in selecting this school was that both achievement test scores and teacher-rating scores were available for students.

**Instruments**

*The Midwest State Achievement Test (MSAT).* MSAT measures how well students have met the state learning standards (Illinois State Board of Education [ISBE], 2009). The test is also aligned with the Stanford Achievement Test (SAT10) and includes some items from the SAT10 to allow for national norm comparisons in addition to measuring the state learning standards. Content covered by the MSAT includes reading and math every year, beginning in Grade 3, as well as science in Grades 4 and 7. Reported alpha reliability values for each grade and content area range from .89 to .94 (ISBE, 2009). In addition, the test manual presents detailed information with regard to item development, including subscale information functions and classical test and item response theory methods of development. Similarly, detailed information is available on MSAT validity, including item- and subscale-total correlations, concurrent validity coefficients with the SAT10 (.91 and more), and details on the three-parameter item response model used for most item revision and evaluation.

*HOPE Scale.* The second instrument used in this study was the HOPE Scale (Peters & Gentry, 2010), an 11-item teacher-rating instrument that measures observable characteristics of Academic and Social dimensions of giftedness. Teachers respond to 11 items using a 6-point frequency scale (1 = *never to*; 6 = *always*). It was designed to help teachers identify students from low-income families for gifted services and developed using two samples totaling approximately 7,700 K-5 students, which yielded valid (Comparative Fit Index [CFI] = .95, Tucker Lewis Index [TLI] = .94, Root Mean Square Error of Approximation [RMSEA] = .10) and reliable ( $\alpha$  reliability estimates: .93 academic and .92 social) information from the data regarding typical gifted and talented behaviors and no differential item functioning when comparing racial/

**Table 2.** Cut Scores and Number of Students From Low and Regular Income Families Identified in Reading, Math, and Science Using Local, General Norms

Content	25%		10%		5%		Total students
	Low SES	Regular SES	Low SES	Regular SES	Low SES	Regular SES	
	<i>n</i>	<i>n</i>	<i>n</i>	<i>n</i>	<i>n</i>	<i>n</i>	
Read08 (Cut-point)	16 (244)	34	7 (259)	13	3 (268)	12	176
Math08 (Cut-point)	13 (257)	34	3 (271)	17	2 (291)	9	179
Sci08 (Cut-point)	5 (235)	17	1 (259)	9	0 (269)	6	82
Read09 (Cut-point)	16 (248)	53	8 (266)	24	4 (272)	18	265
Math09 (Cut-point)	12 (262)	59	4 (283)	27	0 (296)	19	266
Sci09 (Cut-point)	11 (230)	14	6 (247)	5	4 (255)	4	90

Note: Because the distribution of scores is slightly skewed, there are more students at some levels than would be expected. Cut points are in parentheses for each measure at each percentage.

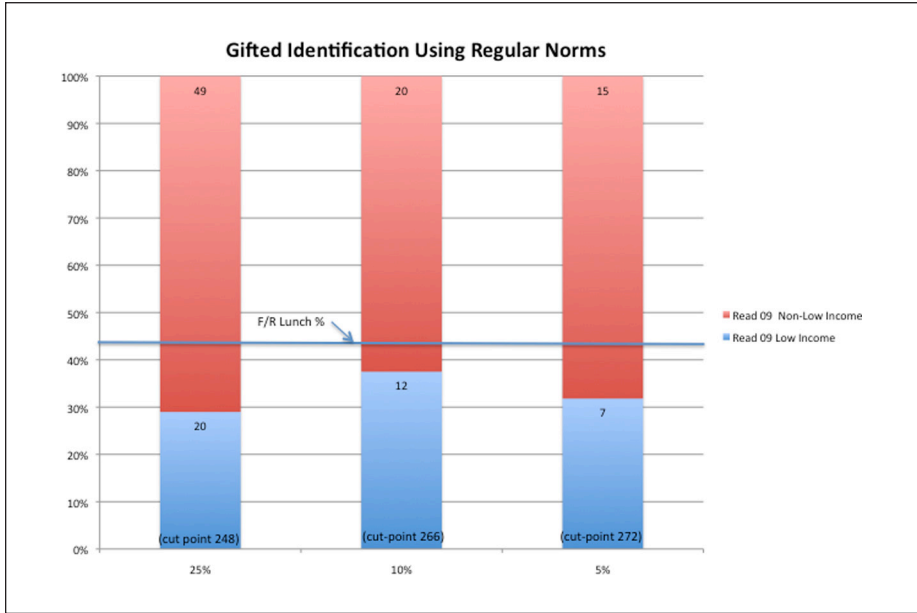
ethnic or income groups (Peters & Gentry, 2010). An example item from the Social scale is “Is sensitive to larger or deeper issues of human concern,” whereas an example from Academic scale item is “Has desire to work with advanced concepts and materials.”

## Data Analysis and Results

### Group-Specific Norms

Three scenarios, in which 5%, 10%, or 25% of students would be identified for gifted and talented services based on MSAT subscale scores, were created and compared to address the research question. At this point, it is important to note that we neither condone nor recommend this practice. However, we use it here as a typical practice in gifted and talented student identification. Table 2 presents the applicable cut scores at three typical percentages often used to classify students as gifted and talented in reading (i.e., top 5%, top 10%, and top 25%). Test scores are presented over a 2-year period, meaning two separate testing/identification occasions. The numbers in parentheses indicate the overall number of students who would be identified based on each particular cut score. This information is presented graphically in Figure 1 using Reading 2009 data for the example.





**Figure 1.** Proportions of students identified as gifted by income group using general norms. Note. The “F/R Lunch %” line indicates the level of proportional representation.

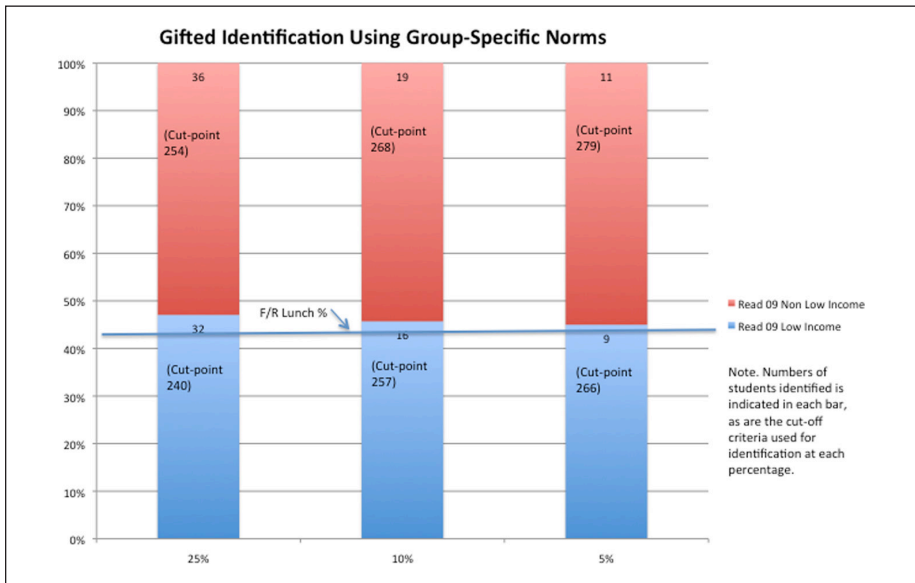
The number of total students from 2008 to 2009 testing occasions differed mostly because 2009 included an additional grade level (those who were in second grade in 2008 were now eligible for the MSAT). The results in Table 2, especially those in the 5% column, represent a fairly common “gifted program” with regard to homogeneity for each content area. The separation of identification by domain of talent is also consistent with current thinking and suggested practice (Lohman & Renzulli, 2007). At this point, no distinction was made for any subgroup membership meaning that the same cut scores were used for all students regardless of their socioeconomic background. Figure 1 graphically depicts the proportional representation of the identified gifted population for each of the three cut scores. The horizontal line indicates proportional representation.

Table 3 separated students for whom information was available into those who were and were not eligible for the free or reduced lunch program. From these data, the relationship between family income and academic achievement is clear. The cut scores at each percentage range from small differences or even being equal in the 2009 Science test, to as different as 31 points in the 2008 Science test. This breakdown presents an example of what frequently happens in gifted and talented programs. Students from low-income families are less-frequently identified as gifted and talented because their average achievement test scores are lower than those of their higher income peers. If the general norms presented in Table 2 and Figure 1 at the top 10%

**Table 3.** Numbers of Students Identified Using Set Percentages of Top Scoring Students and Group-Specific Normed Cut Scores

Content	Low-income				Regular income			
	MSAT score ( <i>n</i> identified)			Total students	MSAT score ( <i>n</i> identified)			Total students
	25%	10%	5%		25%	10%	5%	
Read08	26 (237)	15 (247)	7 (259)	92	29 (247)	12 (268)	7 (274)	84
Math08	25 (243)	12 (258)	7 (267)	94	24 (264)	11 (285)	7 (299)	85
Sci08	12 (218)	5 (232)	3 (244)	41	11 (251)	6 (269)	3 (275)	41
Read09	32 (240)	16 (257)	9 (266)	125	36 (254)	19 (268)	11 (279)	140
Math09	34 (254)	16 (269)	8 (277)	125	36 (267)	20 (289)	10 (302)	141
Sci09	14 (227)	6 (247)	4 (255)	47	13 (235)	5 (251)	5 (255)	43

Abbreviation: MSAT = Midwest State Achievement Test. Because the distribution of scores is slightly skewed, there are more students at some levels than would be expected. Cut Scores are in parentheses for each subject and each percentage.



**Figure 2.** Proportions of students identified as gifted by income group using group-specific norms

level were used (and this practice is common), even the low-income students scoring in the top 5% of the group-specific norm in the 2008 Math test would fail to be identified. In this case, the general norm at the 10% level of 271 is higher than even that of the top 5% level for the low-income group-specific norm.

Figure 2 includes the 2009 Reading data (from Table 3) and depicts what would happen if group-specific norms were used to make identification or placement decisions.

As shown in Figure 2, using group-specific norms for students from low-income families results in almost perfect proportional representation at any percentage level. When the income-group-specific norms for reading at the top 25% level were used, approximately 47% of those students identified would be from low-income families. As 43% of the school's students are eligible for the free or reduced lunch program, this is a much more proportional outcome than is shown in Figure 1 in which only 22% of the "identified" students come from low-income families. It is important to note that simple proportionality is not the only goal; rather, the most important outcome is to use identification to better foster the talents of all students.

The practice of using cut scores to determine program eligibility is common (Barton & Coley, 2009). For example, if school administrators were to identify those students as gifted who score in the top 25% of their local school norm (grade-level norms could also be used), then the scores in the first column of Table 2 would be the cut scores with the number in the parentheses indicating how many students would be identified in each content area. If the scores in the first column of Table 2 are compared with the scores presented in Table 3 for students from low-income families, the problem becomes clear. Using full-school local norms will still identify few low-income students. In this case, the scores for students in the top 25% of the school are comparable with the scores of students from low-income families who scored in the top 5% (except for 2009 Math data in which only students in the top 2% would be identified). As 43% of students in this school are eligible for the federal free or reduced lunch program, following a policy of a general cut score at the school-level would lead to underrepresentation of students from low-income families.

### *Combining Local Norms and Teacher-Rating Scales*

Using local group-specific achievement tests norms with this example data set resulted in more proportional numbers of low-income students being identified as gifted. When a teacher-rating scale is used in conjunction with local norms, additional students, including those who underachieve or who perform poorly on achievement tests, can also be located. Included would be students who achieve at levels lower than their higher income peers, but at high achievement levels when compared with their specific income group. These students have been recognized by their respective teachers as having demonstrated characteristics commonly associated with giftedness and talent in one or more domains. However, for some students, these skills translate to high MSAT scores. Lohman and Renzulli (2007) noted that "adding points [for teacher rating scales] to points earned for test scores can increase the diversity of the population of students who are admitted [to a gifted program], especially when [teacher] ratings are weighted heavily" (p. 2).

**Table 4.** Example Process for Combining MSAT and HOPE Scale Scores for Placement Purposes

MSAT scores (in any content area)	HOPE Scale scores (academic subscale)		
	Bottom quartile ratings	Middle 50% ratings	Top quartile ratings
All students in top 10% of general norm	III	II	I
Low-income students in top 10% of group-specific norm	VI	V	IV
Low-income students in top 25% of group-specific norm	IX	VIII	VII
Alternate pathway			X

Abbreviation: MSAT = Midwest State Achievement Test. Method adapted from Lohman and Renzulli (2007).

Combining the two practices of group-specific norms and teacher-rating scales can follow a modified version of the process described by Lohman and Renzulli (2007). Although originally designed using the Cognitive Abilities Test (CogAT; Lohman & Hagen, 2001) and the Scales for Rating the Behavioral Characteristics of Superior Students (SRBCSS; Renzulli et al., 2002), this process can be applied using different instruments. Table 4 presents the combination of the MSAT and HOPE Scale scores used to make identification or placement decisions.

Applying this method yielded 10 different groups depending on students' scores on the MSAT and whether or not they received a HOPE Scale score in the top quartile, the middle two quartiles, or the bottom quartile. Because teachers are instructed to complete the HOPE Scale by rating each student as compared with other students of similar background or experience, a general local norm can be used. For the most part, this method does not separate students into gifted/nongifted groups. Instead, in this example, the decision was already made that students scoring at or more than certain percentiles were advanced to the point that they require additional challenge or programming. Importantly, the practitioner needs to decide, after consulting state or local regulations and definitions, as well as the local school curriculum, which groups of students are in need of additional services beyond what is provided in the general education setting (see U.S. DOE, 1993 definition in an earlier section). Thus, all students who fall into any of the cells could potentially be labeled as "gifted" or more appropriately could be identified for special services to help them grow based on their individual needs and current achievement level. Table 5 depicts the application of Reading 2009 MSAT data to this process.

Taking all the students who scored at or more than their given norm and placing them in a single program as if they all had the same learning needs would be inappropriate. For example, students who received top quartile versus middle half HOPE Scale academic scores could have very different scores on the Reading subscale of the MSAT. In fact, according to the 2010 MSAT score ranges chart (ISBE, 2010), a score

of 257 is equivalent to an “exceeds expectations” level score for a sixth-grade student; whereas, a score of 268 is equivalent to an “exceeds expectations” score for a seventh-grade student. In grade-level terms, these two groups are, on average, a full grade level apart in terms of their achievement. As all the students in the present sample were in Grades 4 through 6 for the 2009 MSAT occasion, they are all well above the average. However, it is also clear that the learning needs of students in these two groups would be very different regardless of their chronological grade level. Some are working at the level typical of a sixth-grade student, whereas others have mastered material typical of an advanced seventh-grade student. For this reason, students in Cell I might need to be subject accelerated; whereas, the students in Cell III, due to their high test scores but low ratings, could require greater depth of work and/or a different educational environment. In addition, simply providing accelerated or advanced content may not be sufficient. Changes in instruction, pedagogy, or even content with regard to examples or class materials may need to be made to best match instruction with student needs and background experiences. Of course, such decisions would depend on the services and levels of challenge provided in the general education setting. This needs-based perspective on educational programming and gifted and talented identification is similar to what is currently being discussed within the Response to Intervention framework (Rollins, Mursky, Shah-Coltrane, & Johnsen, 2009). All students in Cells I, II, and III would be easily identified, as they have the highest test scores. More information is needed for the student in Cell III, as this student has high scores but low teacher ratings. Perhaps, this was a case of conflicting personalities between teacher and student, low student motivation, or simply a student who did not impress his or her teacher despite having high ability.

Combining the two measures (achievement tests and teacher-rating scores) accomplished two things. First, it broke down the students scoring in the top 10% into those who received high teacher ratings and those who did not. This kind of distinction is important when it comes to programming as those students who do not receive high ratings could be more likely to underachieve, go unrecognized by their teachers as gifted, have trouble learning in a typical classroom environment, or fall behind. As indicated previously, students in Cells I, II, and III could all be identified with the understanding that there may be differences in the services they need. Similarly, students in Cells IV, V, and VI would be identified because they have achieved at a high level relative to their peer group. However, they would likely still have different needs from those students in Cells I to III as they have mastered different levels of concepts relative to those students in the top three cells (as evidenced by the different cut scores of 266 and 257, respectively). In other words, the content the two groups have mastered and the content they are ready to learn is going to be different (note that as Figure 1 presents, some low-SES students do achieve just as high as their higher-SES peers). Just because two groups of students would now be called “gifted,” it does not mean they have identical learning needs.

The second benefit of using the two measures is that the teacher-rating scale can help find individuals who might be missed by relying on test scores alone, but who received

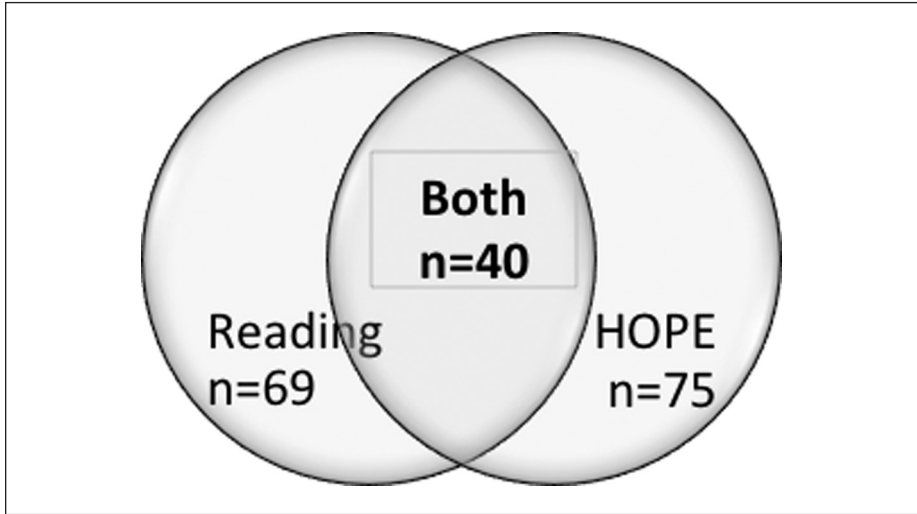
high ratings by a teacher familiar with their classroom performance. These students stood out in some way, despite their less-than-remarkable test scores. Students in Cell IX received achievement scores at or more than the 25th percentile for their income group and also received teacher ratings in the bottom quartile. These students would not likely be identified at that time for advanced programming. These students could be underachieving, disenchanted with school, and/or could be similar to the students described by Plucker et al. (2010) as those who have fallen behind. They are the students who need enriched programs to encourage them and help them develop their skills and talents. Their basic academic needs might be met in the general classroom, but to develop their potential, they will require targeted, thoughtful intervention.

Students in Cells VII and VIII received high and medium teacher ratings and scored within the top 25% of achievement within their income group. These students need to be observed and attention needs to be directed to determine whether their educational needs are being met, whether they are ready for additional programming, or whether more information is needed. These are also potentially underachieving students who need an appropriate educational environment to fully develop their potentials. Students in Groups VII and VIII may also be poor test takers, meaning that additional assessment or subjective information may be needed to determine proper educational placements or interventions.

**Table 5.** Application of Lohman and Renzulli (2007) Method to MSAT and HOPE Scale Scores

	HOPE Scale scores (academic subscale)		
	Bottom quartile ratings	Middle 50% ratings	Top quartile ratings
MSAT scores (reading)	Academic <16	Academic 16-24	Academic >24
All students in top 10% of general norm: cut score = 266 ( <i>n</i> = 32)	III 1	II 10	I 21
Low-income students in top 10% of group-specific norm: cut score = 257 ( <i>n</i> = 16)	VI 2	V 6	IV 8
Low-income students in top 25% of group-specific norm: cut score = 240 ( <i>n</i> = 31)	IX 5	VIII 12	VII 14
All students who did not score in the top 25% of group-specific norm (240 low-income) (248 general)	Not identified	Not identified	X 11 low-income 24 high-income

Abbreviation: MSAT = Midwest State Achievement Test. The HOPE Scale Score columns indicate above-average scores on either subscale.



**Figure 3.** Overlap between HOPE Scale academic and MSAT reading 2009 scores

One issue with the system presented in Table 5 is that Cells I to IX required students achieve at or above the 75th percentile. This is a common but questionable practice as was discussed in the literature review. Because this system still places the achievement test as the required gate in the path toward identification, an alternative pathway was needed to locate those students who received especially high teacher ratings (top quartile) but who did not make it into the top 25% of achievement test scorers. This final cell (X) is represented as the alternative pathway. Simply put, because of the imperfections in educational assessment, alternate pathways are important. Achievement tests have limitations. Students could have a bad test day or suffer from test anxiety, and a variety of reasons and errors exist that could account for an observed score that does not accurately reflect the achievement level of the child (true score).

Cell X contains students who received teacher ratings in the top 25%, but who also received achievement test scores in the bottom 75%. Including this cell and this alternate pathway located an additional 35 students. These students are represented on the right side of the Venn diagram presented in Figure 3. The overall correlation between the HOPE Scale academic subscale and the MSAT Reading 2009 scores was .55.

These could be talent pool students or those who might enter programming on a trial basis to determine whether the program helps them achieve or whether some other kind of programming is necessary. It is important to note that their teachers saw something in them even in the absence of high test scores.

### Recommendations

We must make the important point that we do not recommend the practice of using a state-level achievement test as the sole criterion for gifted program placement, as

these tests are often plagued with validity issues especially when it comes to assessing students scoring in the highest and lowest ranges. Many achievement tests can be included in a larger identification protocol, but they should only be included alongside other measures and information, particularly those that have been shown to assess the student needs that the resulting program is meant to serve. Many schools still rely solely on achievement or aptitude tests for program identification, even though this practice is neither ideal nor recommended. It is important to note that using group-specific norms can, as shown in our example, enhance the use of achievement measures to yield more proportional representation of underserved students in gifted programs.

It is clear from the results discussed earlier that using group-specific norms for low-income students will locate those students who have demonstrated high achievement (when compared with their peers), but who also often go unnoticed. This practice results in a more proportional representation of students from low-income families in the identified gifted population of a school. It is also clear that combining such a practice with the use of a teacher-rating instrument will locate even more students than by using group-specific norms alone. However, placing both groups of students, those who meet the general group and those who meet the group-specific norms, in a typical gifted program is unlikely to provide the services needed for many students. Instead, multiple levels of service must be applied to allow for all students to grow and develop within the various content areas. In addition, if students from low-income families (especially those in Cells VIII and IX) are to be successful now or in the future in traditional academic gifted programs, additional support structures and/or more culturally relevant content (Castellano & Frazier, 2010) may be required. Talent development programs and gifted services must be responsive to the strengths and needs of the students they serve, not simply a program of more, harder work for students from parents with wealth and social capital.

## **Limitations**

A major caveat underscores all of the results and information presented in this article. This research dealt solely with traditional forms of academic achievement that are typically taught in K-12 schools and measured by a state-level achievement test. This does not mean these are the only forms of giftedness or talent, just that these were used as a convenient and common example. Another limitation is that this research relied on a state-level achievement tests. These tests can be limited by low ceilings, which potentially limit the validity of the information yielded with regard to top-performing students. This does not limit the utility of the idea, but could mean low MSAT test validity, especially concerning the top performers. Unfortunately, many schools use state tests as their main source of student achievement measures, which are central to their identification processes (NAGC, 2009).



The methods presented in this study illustrate a single example from one school, using only achievement data and teacher-rating scores. As Lohman and Renzulli (2007) suggested, having an additional aptitude test score could improve this process. In addition, this article used arbitrary percentages to illustrate the example. In fact, we advocate that percentages should be avoided altogether and instead, who constitutes “the gifted” should be based on those students who have needs not currently being met in general education as per the U.S. DOE (1993) definition. However, in this case, the method and procedure effectively identified more students from low-income families and students who may be underachieving. Identifying these students can result in services to help them more fully develop their talents and potentials.

## **Future Research**

As we indicated, very little applied, empirical work has been done in the area of local and group-specific norms as a means of identifying underrepresented students with gifts and talents. This method needs to be researched with regard to underrepresented racial/ethnic groups as well as replicated in larger more representative settings. Questions also need to be asked with regard to the school policy and legal implications that such practices carry. More research is also needed regarding the future success rate of those students identified via alternate pathways. In this case, the HOPE Scale was used because it provided additional information from the teachers’ perspective. We note that although the HOPE Scale has undergone validation studies of content and construct validity, more work remains concerning its long-term predictive validity. In addition, the utility and quality of the identification systems presented here will improve with more use and with more research. If underrepresentation is to be addressed, then bold actions are warranted, as without these actions, the field will continue to see few strides in this area and ever-widening gaps.

## **Declaration of Conflicting Interests**

The authors declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: The two authors of this paper are also the authors of the HOPE Teacher Rating Scale which is available free of charge from <http://www.wix.com/purduegeri/projecthope>.

## **Funding**

The authors disclosed the receipt of the following financial support for the research, authorship, and/or publication of this article: Portions of this research were supported by a Summer Faculty Fellowship grant awarded to the first author by the University of Wisconsin - Whitewater. Project HOPE was initially supported by the Jack Kent Cooke Foundation.

## **Note**

1. Some of the literature review presented here was included in a doctoral dissertation by the first author.

## References

- Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice Hall.
- Barton, P. E., & Coley, R. J. (2009). Measuring the achievement elephant. *Educational Leadership*, 66(4), 30-34.
- Castellano, J. A., & Frazier, A. D. (2010). *Special populations in gifted education*. Waco, TX: Prufrock Press.
- Donovan, M. S., & Cross, C. T. (2002). *Minority students in special and gifted education*. Washington, DC: National Academy Press.
- Gagné, F. (1994). Are teachers really poor talent detectors? Comments on Pegnato and Birch's (1959) study of the effectiveness and efficacy of various identification techniques. *Gifted Child Quarterly*, 38, 124-126.
- High, M. H., & Udall, A. J. (1983). Teacher ratings of students in relation to ethnicity of student and school ethnic balance. *Journal for the Education of the Gifted*, 6, 154-166.
- Hodge, R. D., & Cudmore, L. (1986). The use of teacher-judgment measures in the identification of gifted pupils. *Teaching and Teacher Education*, 2, 181-196.
- Illinois State Board of Education. (2009). *Illinois Standards Achievement Test Technical Manual*. Retrieved from [www.isbe.net/assessment/pdfs/iaa/IAA\\_Tech\\_Manual\\_09.pdf](http://www.isbe.net/assessment/pdfs/iaa/IAA_Tech_Manual_09.pdf)
- Illinois State Board of Education. (2010). *Scale Score Ranges that Define Student Performance Levels on the 2010 ISAT Scales*. Retrieved from [www.isbe.net/assessment/pdfs/cut\\_points\\_10.pdf](http://www.isbe.net/assessment/pdfs/cut_points_10.pdf)
- Lohman, D. F. (2005). An aptitude perspective on talent: Implications for identification of academically gifted minority students. *Journal for the Education of the Gifted*, 28, 333-360.
- Lohman, D. F. (2006). *Identifying academically talented minority students* (Research Monograph No. RM05216). Storrs: The National Research Center on the Gifted and Talented, University of Connecticut.
- Lohman, D. F., & Hagen, E. P. (2001). *Cognitive Abilities Test—Form 6*. Itasca, IL: Riverside Publishing.
- Lohman, D. F., & Renzulli, J. (2007). *A simple procedure for combining ability test scores, achievement test scores, and teacher ratings to identify academically talented children*. Unpublished manuscript.
- Matthews, M. S. (2007). Assessing the impact of a proposed rule change on the identification of gifted English language learners. *Gifted Children: An Electronic Journal of the AERA SIG—Research on Giftedness, Creativity, and Talent*, 2(1), 12-13. Retrieved from <http://docs.lib.purdue.edu/giftedchildren/vol2/iss1/5>
- Matthews, M. S., & Kirsch, L. (2011). Evaluating gifted identification practice: Aptitude testing and linguistically diverse learners. *Journal of Applied School Psychology*, 27, 155-180. doi:10.1080/15377903.2011.565281
- Miller, L. S. (2004). *Promoting sustained growth in the representation of African Americans, Latinos, and Native Americans among top students in the United States at all levels of*

- the education system* (Research Monograph No. RM04190). Storrs: The National Research Center on the Gifted and Talented, University of Connecticut.
- National Association for Gifted Children. (2009). *State of the states in gifted education*. Washington, DC: Author.
- National Center for Educational Statistics. (n.d.). *Number and percentage of public school students eligible for free or reduced-price lunch, by state: 2000–01, 2005–06, 2006–07, and 2007–08*. Retrieved from [http://nces.ed.gov/programs/digest/d09/tables/dt09\\_042.asp](http://nces.ed.gov/programs/digest/d09/tables/dt09_042.asp)
- Pegnato, C. W., & Birch, J. W. (1959). Locating gifted children in junior high schools: Comparison of methods. *Exceptional Children, 25*, 300-304.
- Peters, S. J., & Gentry, M. (2010). Multi-group construct validity evidence of the HOPE Scale: Instrumentation to identify low-income elementary students for gifted programs. *Gifted Child Quarterly, 54*, 298-313.
- Peterson, J. S. (1999). Gifted—through whose cultural lens? An application of the postpositivistic mode of inquiry. *Journal for the Education of the Gifted, 22*, 354-383.
- Peterson, J. S., & Margolin, L. (1997). Naming gifted children: An examination of unintended reproduction. *Journal for the Education of the Gifted, 21*, 82-100.
- Pfeiffer, S. I., & Jarosewich, T. (2003). *Gifted Rating Scales*. San Antonio, TX: Pearson.
- Plucker, J. A., Burroughs, N., & Song, R. (2010). *Mind the (other) gap!: The growing excellence gap in K-12 education*. Bloomington, IN: Center for Evaluation and Education Policy.
- Reis, S. M., & McCoach, D. B. (2000). The underachievement of gifted students: What do we know and where do we go? *Gifted Child Quarterly, 44*, 152-170.
- Renzulli, J. S., Smith, L. H., White, A. J., Callahan, C. M., Hartman, R. K., & Westberg, K. (2002). *Scales for rating the behavioral characteristics of superior students*. Mansfield Center, CT: Creative Learning Press.
- Rollins, K., Mursky, C., Shah-Coltrane, S. V., & Johnsen, S. K. (2009). RtI models for gifted students. *Gifted Child Today, 32*(3), 21-30.
- Ryser, G. R., & McConnell, K. (2004). *Scales for identifying gifted students*. Waco, TX: Prufrock Press.
- Stambaugh, T. (2007). Next steps: An impetus for future directions in research, policy, and practice for low-income promising learners. In J. VanTassel-Baska & T. Stambaugh (Eds.), *Overlooked gems: A national perspective on low-income promising learners* (pp. 83-88). Washington, DC: National Association for Gifted Children.
- Swanson, J. D. (2006). Breaking through assumptions about low-income, minority gifted students. *Gifted Child Quarterly, 50*, 11-25.
- United States Department of Education. (1993). *National excellence: A case for developing America's Talent*. Washington, DC: Author.
- Valencia, R. R., & Suzuki, L. A. (2001). *Intelligence testing and minority students*. Thousand Oaks, CA: SAGE.
- VanTassel-Baska, J. L. (2008). Epilogue: What do we know about identifying and assessing the learning of gifted students? In J. L. VanTassel-Baska (Ed.), *Alternative assessment with gifted and talented student* (pp. 309-319). Waco, TX: Prufrock Press.

- Viadero, D. (2006, November 8). Scholars test out new yardstick of school poverty. *Education Week*, 26(11), 1, 15.
- Wyner, J. S., Bridgeland, J. M., & Diulio, J. J. (2009). *The achievement trap* (Rev. ed.). Washington DC: The Jack Kent Cooke Foundation.
- Yoon, S., & Gentry, M. (2009). Racial and ethnic representation in gifted programs: Current status of and implications for gifted Asian American students. *Gifted Child Quarterly*, 53, 121-136.

Copyright of Journal of Advanced Academics is the property of Sage Publications Inc. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.