Check for updates

AERA Open
April-June 2019, Vol. 5, No. 2, pp. 1–18
DOI: 10.1177/2332858419848446
Article reuse guidelines: sagepub.com/journals-permissions
© The Author(s) 2019. http://journals.sagepub.com/home/ero

Effect of Local Norms on Racial and Ethnic Representation in Gifted Education

Scott J. Peters
University of Wisconsin–Whitewater

Karen Rambo-Hernandez
West Virginia University

Matthew C. Makel
Duke University

Michael S. Matthews
University of North Carolina–Charlotte

Jonathan A. Plucker
Johns Hopkins University

Educators have sought to understand and address the disproportional representation of students from certain student subgroups in gifted education. Most gifted identification decisions are made with national comparisons where students must score above a certain percentage of test takers. However, this approach is not always consistent with the overall goal of gifted education. Scholars have long argued for the use of local normative criteria to increase the diversity of students identified for gifted services, and although some districts across the country have applied such recommendations, little research has been carried out. In this study, we use a large data set to assess the extent to which identifying gifted students with either school-level norms or a combination of national and school-level norms would improve gifted education representation rates for students who are from African American and Latinx families. A preprint of this registered report and this project's preregistration documentation are available at https://osf.io/z2egy/.

Keywords: *assessment, descriptive analysis, diversity, equity, gifted education, hierarchical linear modeling, policy analysis*

Disproportionality rates have been reported along racial/ethnic lines in areas including school discipline, average test scores, hiring practices, college enrollment, and a host of other important student outcomes (e.g., Cruz & Rodl, 2018; Gupta-Kagan, 2017; U.S. Department of Education, 2016; Williams, Bryant-Mallory, Coleman, Gotel, & Hall, 2017). Furthermore, disproportionality is evident in rates of enrollment in K–12 gifted education programs by certain racial, ethnic, income, language, and disability subgroups (e.g., Peters, Gentry, Whiting, & McBee, 2019; Yoon & Gentry, 2009), and gifted education has been challenged on the basis of equity and corresponding concerns about whether its practices exacerbate inequality across student subgroups (e.g., Garland, 2013). Much of this criticism arises from the observation that students served by gifted education programs tend to be from European American, Asian American, or upper-income backgrounds—an observation that has been documented since at least the 1970s (Peters et al., 2019; Yoon & Gentry, 2009).

Not surprisingly, gifted identification disparities occur alongside large differences in rates of advanced performance among subgroups of students, a phenomenon known as *excellence gaps* (Plucker, Hardesty, & Burroughs, 2013). For example, on the 2017 National Assessment of Educational Progress (NAEP) mathematics assessment for Grade 4 students, 2% of African American and 3% of Hispanic[1] students scored at the advanced level, whereas 24% of Asian American and 11% of European American students[2] scored in the

Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (http://www.creativecommons.org/licenses/by-nc/4.0/) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).

advanced range. These stark differences in advanced performance will have far-reaching cultural and economic implications if they remain unaddressed, because the subgroups less frequently performing at advanced levels now represent well over half of the U.S. student population (Plucker & Peters, 2018).

Although the causal mechanisms behind excellence gaps have yet to be explored, Plucker and Peters (2018) suggested that disproportional access to advanced educational services vis-à-vis disproportionality in gifted identification is one of the drivers. The presence of disproportionality suggests that many students who remain unidentified would benefit from placement into gifted education programming. In this article, we explore one potential route to shrinking such disproportionality in gifted program participation: the use of local building-level norms.

Quantifying Disproportionality

Disproportional representation has been quantified in three related ways, each with its own strengths and weaknesses: aggregate numbers, enrollment relative to base rate, and conditional probability of identification. The first approach expresses disproportionality in terms of aggregate numbers of identified students. For example, on the 2017 NAEP mathematics assessment for fourth graders, African American students showed a mean score of 223, as opposed to a mean score of 248 for European American students. This difference is almost a full standard deviation. If students were identified as gifted on the basis of a score on this or similar measures of academic achievement, fewer African American students would be identified than European American students—thereby resulting in aggregate racial disparities. Although measuring underrepresentation in this manner is appealing in its simplicity, it fails to account for the proportionality of each group within the larger student population (i.e., the base rate). In other words, one would expect to find smaller numbers of African American students identified as gifted because African Americans also constitute a smaller percentage of the overall student population, but this is not clear from the raw numbers alone.

Enrollment relative to base rate, often called a *representation index*, offers an improvement over aggregate numbers by reporting the proportion of students identified from each group relative to their proportion in the overall student population. A representation index is a form of the general relative risk calculation and is the ratio between any given group's representation in the identified gifted population and its representation in the overall student population. This is the approach most frequently used in education research, including studies of disproportionality in other areas of education, such as special education services (e.g., Morgan, Farkas, Hillemeier, & Maczuga, 2017) and school discipline (e.g., Gregory, Cornell, & Fan, 2011). For example, Peters et al. (2019) found that African American and Latinx students were represented in identified gifted populations at approximately 57% and 70% of these students' prevalence in the overall K–12 student population. At the same time, students who self-identified as Asian American or European American were 201% or 118% as represented in identified gifted populations, respectively, as in the overall K–12 population.

A third way to operationalize disproportionality is to look at the probability of a student being identified after controlling for relevant background factors (conditional probability of identification). This method has received increased attention in the field of special education (see Morgan et al., 2017) because it better distinguishes *disproportional representation* from *underrepresentation*; the latter term implies what a group's representation actually should be, which raw numbers or even relative percentages cannot fully address. As compared with the previous two methods, the identification probability approach evaluates disproportionality while attempting to control for background factors known to be relevant, thereby allowing a determination of whether and to what extent a student's ethnicity or other characteristic taken in isolation may drive disproportionality.

Using the identification probability approach, Siegle, McCoach, Gubbins, Long, and Hamilton (2018) found that even after controlling for third-grade reading achievement, mathematics achievement, student demographics, school and district socioeconomic status, school and district achievement, and the percentage of students identified as gifted in the district and school, students from African American, Latinx, or low-income families remained less likely to be identified for gifted education services. Grissom and Redding (2016) found similar results but with additional nuance: for Hispanic students, the gap in probability of identification was fully explained after controlling for student background factors, such as prior achievement and family income. The same could not be said for the gap in identification probability for African American students, for whom the race of the teacher was also a contributing factor. These studies' findings suggest that it is not simply lower group mean scores that prevent underrepresented students from being identified but that additional factors also influence a student's probability of being identified as gifted (see also Hamilton et al., 2018).

The benefit to considering disproportionality through the lens of identification probability is that this approach controls for other relevant background factors, thus clarifying the source of the disproportionality. Observed mean score differences on standardized tests can be included in a model that allows race or ethnicity to be examined in isolation from other potentially confounding variables. Although the identification probability approach is important from a basic science perspective, it may be less helpful from a policy

perspective because applying such conditional identification methods in schools would be complex; hence, both perspectives are useful—identification probability and enrollment relative to base rate.

Gifted Identification Policy and Practice

The difference between gifted education and other areas of exceptional student education is that the procedures for deciding which students are served in gifted education vary widely across and within states. The number of students identified as gifted depends largely on policies developed at the state and local levels, and these vary widely across different gifted education models as well as in actual practice (Callahan, Moon, & Oh, 2014). The National Association for Gifted Children has suggested the overall prevalence of gifted students as including the top 10% or less within a given domain. Some states (e.g., Arizona) mandate a fixed percentage, such as 3% based on a national norm. Renzulli's three-ring model (1978, 2005), used in many schools across the United States, suggests that roughly 15% of students should be identified for gifted services.

Criteria for identifying a visual impairment or a learning disability are relatively consistent across settings due to their basis in federal law, but this is not the case for the identification of a student as gifted. A survey by the National Research Center on the Gifted and Talented (Callahan et al., 2014) highlighted just how widely gifted identification practice and outcomes vary. These authors found that across settings at the elementary level nationwide, the percentage of students identified as gifted ranged from zero to 50%. This extreme variability is due in part to variability in school populations but also to the processes by which students are selected, which vary widely by location. Some states mandate a strict IQ score–based process (e.g., Florida, New Mexico), while others (e.g., North Carolina) delegate many aspects of the process to the local school district. Research by Carman, Walther, and Bartsch (2018) and by Peters and Gentry (2012) adopted a range of cut points as proxies for gifted identification rates, including the top 5%, 10%, or even 25%. Because of the lack of consensus regarding the actual percentage of the population that should be labeled gifted, for the purposes of this article we chose to model two gifted identification rates (5% and 15%) to reflect the range of rates found in actual gifted education settings across the United States.

Cut scores with national norms. A common identification practice across many gifted education settings is the use of cut scores based on national performance metrics. A national norm is most often applied, as in Arizona: "School districts . . . shall identify as gifted at least those pupils who score at or above the ninety-seventh percentile, based on national norms, on a test adopted by the state board of education" (Arizona State Legislature, n.d., 1A). Although academic achievement, ability, and aptitude (including intelligence tests) are the tools most widely used for gifted identification (National Association for Gifted Children, 2015), details of how these tools are used are often less clear; Arizona is an exception in its explicit reference to national norms. Georgia too refers to national norms in its state-mandated gifted identification policy: "Evidence of student performance on a nationally normed standardized test of mental ability, achievement, and creativity" (Georgia Department of Education, 2017, p. 4). However, any "nationally normed" standardized test can be used to collect student performance data, and raw scores can be evaluated with a range of scoring norms, making the actual prevalence of national norm usage unknown.

Limitations of national norm comparisons. Most state and district policies do not specify the norm group to be used when identification decisions are being made. However, those that reference specific scores or percentiles, such as Arizona and Georgia, reference nationally normed instruments. This suggests that national norms are the standard reference group for norm-referenced identification criteria. For example, Tennessee refers to students scoring at or above the 94th percentile—presumably on a national norm (State of Tennessee, 2017). The ubiquity of national norms may be due to convenience, as normative studies with nationally representative samples are typically part of any instrument's development. However, there are at least two problems inherent in this use of national norms. First, students are not randomly assigned to schools; rather, attendance is based largely on residence in local neighborhoods that, in turn, are often highly segregated by income, ethnicity, and other differences. Thus, across schools, national norm comparisons yield drastically different numbers of students identified as gifted. For example, if half of an extremely high-performing school is performing at or above the 95th percentile on a national norm and the 95th percentile is set as the criterion, then half of that school's population would be labeled gifted. With this same cutoff, other schools whose populations are lower achieving overall would identify zero students as gifted.

Second, relying on national norms stands in potential conflict with the current federal definition of giftedness: "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*" (U.S. Department of Education, 1993, p. 3, emphasis added). National norms offer a uniform standard that appears to promote fairness. Typically, they are age based, but otherwise, the extent to which they address student experience or environment is unclear. Two students in the same classroom who grew up as neighbors may have had vastly different educational opportunities, none of which

would be captured by comparing their performance with a national norm. These challenges have led some scholars (Lohman, 2005, 2009; Peters & Gentry, 2012; Plucker & Peters, 2016; Worrell, 2018), advocacy groups (Yaluma & Tyner, 2018), major professional organizations (American Educational Research Association, American Psychological Association, & National Council for Measurement in Education, 2014), and state policies (e.g., Illinois, New Jersey) to call for the use of building-level norms when test score data are used to make gifted placement decisions. Colorado also has endorsed the use of local norms if the school district "determines that such data will enhance services to student groups who may in the future qualify for gifted identification under national norms and/or performance demonstrations" (Colorado Department of Education, 2016, p. 12).

Case for building-level local norms. In this article, we use "local norms" to refer to ranked performance within the school building. This means that the reference group for the gifted identification process is the student's same-grade peers within a given building. Instead of different schools in the same district (or state) having a different proportion of students identified, every school using local norms and a common cutoff would have the same proportion of students identified to receive gifted program services. If the cut score is the top 5% of each building, then each building will always identify 5% of its students as gifted. The logic behind this approach is that these are the students most likely to go underchallenged and thus in need of additional services to be appropriately challenged. From an administrative point of view, identifying consistent numbers of students within schools also simplifies instructional planning: staff allocation is more predictable because the number of students served does not vary as widely across buildings or from one year to the next as when national norms are used to identify learners for gifted services.

The philosophical argument in favor of building-level norms is that within-building peers are a better proxy for experience and environment than are all same-age students from across the country (Peters & Engerrand, 2016). It is at the local building level that most gifted education services are delivered; therefore, the building-level norm is likely the approach most consistent with the intent of the federal definition of giftedness. Furthermore, the purpose of gifted education is to provide identified students with opportunities to be appropriately challenged in their zone of proximal development (Peters, Rambo-Hernandez, Makel, Matthews, & Plucker, 2017) or, as Stanley (2000) put it, to have students learn "only what they don't already know" (p. 216). From this perspective, the role of gifted identification is to place these students into services that are necessary to meet their particular learning needs. Local norms are better suited than national norms to finding the students who are most likely to be underchallenged in their current learning environment.

Implications of local norm comparisons. There are two important implications to using building norms. First, they likely result in varying levels of content mastery being needed to qualify for services depending on which school a child attends, even within the same school district (Carmen et al., 2018), thus making implementation potentially difficult. Second, they may not be as closely connected to broader external metrics, such as "grade level" or "college readiness" measures. The former issue is probably inevitable, although it simply reflects the wide variation in performance levels that already exists across schools, while the latter issue can easily be addressed by retaining national norms for any such comparisons.

Combining multiple criteria. Many gifted identification processes require that multiple criteria be met, often in the form of multiple test scores exceeding certain criteria. The manner in which these criteria are combined can have a strong influence not just on who is identified but also on how many students are identified as gifted (Lakin, 2018; McBee, Peters, & Miller, 2016; McBee, Peters, & Waterman, 2014). To make the eligibility decision, multiple criteria can be combined by using *and* rules (e.g., students need both Criterion 1 and Criterion 2), by using *or* rules (e.g., students need Criterion 1 or Criterion 2), or by using a *mean* rule (e.g., averages of criteria are used). Any of these combination rules can be used with any norm type. Using the *or* combination rule for national and building norms could serve as a compromise between these two disparate approaches. Under this approach, students would be identified if they met either the national norm criterion or the building norm criterion (e.g., top 5% in the nation *or* top 5% in the building). Such a policy would remove any decrease in the number of identified students at overall high-achieving schools while placing a floor on the number of identified students at lower-achieving schools, thus taking advantage of the strengths of each approach.

Combining these approaches has not systematically been considered in the literature, nor has the diversity of the resultant identified population ever been evaluated. The aim of this article is to examine the outcomes of these approaches by modeling them with real data.

Hypotheses

To evaluate the potential of building-level norms to increase the diversity of identified gifted populations, we proposed the following general hypothesis: The more proximate the normative group used for gifted identification decisions, the more racially and ethnically representative the identified population of gifted students will be. Specifically,

we operationalized the general hypothesis into these testable hypotheses:

Hypothesis 1: Using building norms will yield an identified gifted population most representative of the racial/ethnic makeup of the larger K–12 population. Specifically, we hypothesize at least a 20% improvement in the representation index for African American and Latinx students when building norms are used versus national norms.

Hypothesis 2: Using a combination of national plus building norms with the *or* rule (students can qualify as gifted via either building or national norms) will result in a gifted-identified population more representative of the racial/ethnic makeup of the larger K–12 population than national norms only but not as representative as building norms alone.

Hypothesis 3: After controlling for school-level variables, European American and Asian American students will show a higher probability of being identified for gifted education services than African American or Latinx students when national norms are used.

Hypothesis 4: After controlling for school-level variables, African American and Latinx students will have a higher probability of being identified for gifted education services when building norms are used as compared with national, state, or district-level norms.

Methods

Data

Our data came from schools that administered the Northwest Evaluation Association (NWEA) Measures of Academic Progress (MAP) test. We obtained data from NWEA for all participating schools in 10 states—California, Colorado, Illinois, Kentucky, Minnesota, Michigan, Ohio, South Carolina, Washington, and Wisconsin—because these states had the 10 largest percentages of their overall student populations taking the MAP test. The data do not include any identifiable information. Multiple institutional review boards deemed this research not subject to review.

The 10 state data sets included all students in Grades 3 through 8 who took the MAP across a 10-year period: 2007–2008 to 2016–2017. However, because third grade is the most common point for students to be screened for gifted services (Siegle et al., 2018), we decided to analyze only the data from third-grade students who took reading and mathematics MAP assessments in the fall for each of these academic years, for a total of 10 cohorts of third-grade students. We decided to include all schools regardless of their type (e.g., public, private, charter) as long as there were more than five students (on average) tested in third grade at the school.

Measure

The MAP is a computer adaptive assessment of achievement in reading and mathematics. Because MAP is designed for students in Grades K–11 and is computer adaptive, there is little threat of ceiling effects for third-grade students (McCall, Kingsbury, & Olson, 2004), making it ideal for this study. Scores on the MAP demonstrated marginal reliability estimates ranging from .93 to .95 (NWEA, 2011). Concurrent validity estimates of the MAP with state achievement tests have hovered around $r = .80$ for objectively scored items (NWEA, 2011). The year-to-year scaling of the MAP and its measured constructs have been extremely stable (Kingsbury & Wise, 2011; Wang, McCall, Jiao, & Harris, 2012).

Student-level variables. Each student was identified in the data as belonging primarily to one of the following eight races or ethnicities: American Indian or Alaskan, Asian or Pacific Islander (Asian American), Black (African American), Hispanic (Latinx), multiethnic, Native Hawaiian or other Pacific Islander, not specified or other, or White (European American). NWEA-specific labels are those outside the parentheses. Because of the smaller sample sizes and these groups not being a primary focus in our hypotheses, we collapsed American Indian or Alaskan, multiethnic, Native Hawaiian or other Pacific Islander, and not specified or other into one category: other. European Americans served as the reference group. Table 1 lists the number of districts, schools, and students disaggregated by race or ethnicity category in the 10 states represented in the NWEA data that we used.

We created dummy codes to represent each of the other four race/ethnicity categories. The dependent variable was whether the student's observed MAP score in reading was greater than or equal to the cut scores for each of the four comparison norms (e.g., identified gifted under national norms = 1). Similarly, we created another set of variables to indicate whether the students were identified as gifted in mathematics.

Building-level variables. For each school, the data sets included school type (public or private) and setting (city, suburb, town, or rural). We used the type codes associated with each school's first observation in the data set. Among all the schools in the data set, approximately 7% changed setting status over time, but none changed school type. We also created a variable for the proportion of African American or Latinx students by summing the number of these students and dividing it by the total number of students in the school. Across the data set, 199 schools (2%) did not have an indication of public or private school status, and 221 (2.2%) were missing setting data. Thus, in the analyses that included these control variables, the sample size was reduced by approximately 2.2%.

TABLE 1
Number of Districts, Schools, and Students by Race/Ethnicity in the States Represented in the NWEA Data Sets

				African American		Asian American		European American		Latinx		Other	
State	Districts, *n*	Schools, *n*	Students, *n*	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%
Mathematics test													
CA	330	969	257,730	12,933	5	18,445	7	65,535	25	93,287	36	67,530	26
CO	247	874	236,855	7,483	3	6,059	3	122,674	52	55,666	24	44,973	19
IL	547	1,781	578,452	71,223	12	31,980	6	249,074	43	117,109	20	109,066	19
KY	175	594	231,333	15,354	7	2,643	1	150,515	65	10,112	4	52,709	23
MI	575	1,279	328,785	76,160	23	11,121	3	183,500	56	20,466	6	37,538	11
MN	564	1,086	433,898	30,344	7	23,217	5	267,855	62	25,039	6	87,443	20
OH	346	911	207,114	40,205	19	5,349	3	128,438	62	10,190	5	22,932	11
SC	131	712	506,669	167,606	33	7,262	1	256,694	51	37,372	7	37,735	7
WA	223	732	226,116	8,358	4	9,949	4	119,951	53	49,926	22	37,932	17
WI	460	1,075	298,099	21,408	7	10,854	4	190,272	64	24,131	8	51,434	17
Total	3,598	10,013	3,305,051	451,074	14	126,879	4	1,734,508	52	443,298	13	549,292	17
Reading test													
CA	258	662	162,219	7,220	4	13,507	8	45,230	28	59,105	36	37,157	23
CO	244	819	231,391	7,432	3	6,065	3	121,861	53	54,975	24	41,058	18
IL	547	1,779	571,720	69,991	12	31,986	6	248,518	43	115,654	20	105,571	18
KY	175	594	225,946	14,828	7	2,653	1	147,317	65	9,874	4	51,274	23
MI	567	1,270	323,757	74,217	23	11,017	3	182,080	56	20,171	6	36,272	11
MN	559	1,080	432,800	30,547	7	23,213	5	270,133	62	25,256	6	83,651	19
OH	363	968	205,184	39,612	19	5,180	3	127,460	62	10,046	5	22,886	11
SC	131	710	496,093	163,285	33	7,202	1	253,408	51	36,860	7	35,338	7
WA	221	729	220,540	8,251	4	10,030	5	116,844	53	48,950	22	36,465	17
WI	460	1,089	293,731	21,213	7	10,562	4	188,033	64	24,050	8	49,873	17
Total	3,525	9,700	3,163,381	436,596	1	121,415	4	1,700,884	54	404,941	13	499,545	16

Note. NWEA = Northwest Evaluation Association.

Analysis

The first two hypotheses approached disproportionality in terms of identification rate relative to the group's base population rate via representation indices (e.g., Peters et al., 2019; Yoon & Gentry, 2009). The last two hypotheses approached the issue of identification while accounting for the school context (e.g., private vs. public, setting) via the resulting odds ratios (ORs). Of note, OR and representation indices are comparable for low-incidence events (<10%) but not for larger incidences (Davies, Crombie, & Tavakoli, 1998), so qualitative interpretations of OR as if they were enrollment relative to base rate are likely to be fair in low-incidence events. Both the enrollment relative to base rate and OR perspectives were needed to understand the full picture of who is likely to be identified for gifted education services under what type of norm.

Hypotheses 1 and 2. To operationalize the application of national reading and mathematics norms, we determined the percentage of each race/ethnicity that would qualify for gifted services using the top 5% and top 15% as cut scores. We treated our 10-state sample as a population and calculated national norms and cut scores based on the full data set. We recalculated the national norm cut score for every year in our data for a total of 10 cut scores, for the top 5% and the top 15%, in mathematics and reading. First, we calculated the mean and standard deviation of the reading MAP scores in third grade for each fall (2007 through 2017). Second, we calculated the score associated with the top 5% ($z = 1.645$) and top 15% ($z = 1.0366$). A check of data skew revealed that the MAP scores were normally distributed for each year in the data set. We then created two variables to indicate whether each student would qualify for gifted services under either national norm cut score (i.e., whether the student's observed score exceeded the national cut score) in reading and mathematics. We then conducted an identical process using state, district, and building norms, answering the question, would a student's score have placed her or him in the top 15% of the state, district, or building?

In addition to using national, state, district, and school building norms, we evaluated the relative proportionality that resulted from using the *or* rule combination of national and building norms. Under these criteria, students are identified as gifted if they reached either cut score—the top 5% or 15% of the nation or within their building.

Next we reported representation indices, which are the percentage of the identified gifted population (under a given criterion or norm) that identifies with a specific race or ethnicity divided by the percentage of that group within the overall data set. We also calculated change in representation index for all racial and ethnic groups of students to further illustrate changes in enrollment relative to base rate in moving from national to local norms, state to local norms, and district to local norms.

Smallest effect of interest. Because no previous research has conducted such an analysis on a broad scale, we had little ground to make a specific prediction about the magnitude of the effect on identification rates of going from national to local norms. However, this does not prohibit us from establishing benchmarks for what size of an effect we believe would be associated with meaningful change.

Currently, African American and Latinx students are represented at rates of .57 and .70 in gifted programs nationally (Peters et al., 2019). In the study by Carman et al. (2018), which examined a single large district, African Americans were closer to 25% and Latinx closer to 50%. For the purposes of this article, we adopted a benchmark of a 20% increase in representation as our smallest effect of interest (Lakens, 2014)—in other words, an increase in representation due to changing from national to building norms ranging from 20% to the point of perfect proportionality (i.e., 1.0 would be considered meaningfully effective). If the more proximal norm group yielded an enrollment relative to a base rate increase ≥20% as compared with its referent, then we asserted that this constituted a "better" identification strategy, especially given its low cost. We acknowledge that this is an arbitrary determination and that others may support a different threshold.

Hypotheses 3 and 4. To address Hypotheses 3 and 4, we built multiple hierarchical generalized linear models using penalized quasi-likelihood estimation in HLM 7.03 (Raudenbush, Bryk, Cheong, Congdon, & du Toit, 2013) and report the unit-specific model results. To address differences in the probabilities of being identified as gifted on the basis of national, state, district, or building norms, we built four-level hierarchical generalized linear statistical models. These models allowed us to determine the probabilities of identification for students of different race/ethnicities at typical schools (e.g., we controlled for school type, setting, and percentage of minority students). We used the conservative recommendation for statistical significance proposed by Benjamin et al. (2018), which suggests that $p < .005$ results are statistically significant and $p < .05$ results are simply suggestive of a trend. For the effect size, we calculated ORs for each comparison of interest using the levels for small, medium, and large effects described by Chen, Cohen, and Chen (2010). We also report the predicted probabilities for each group under the various norming methods. See the Methodological Appendix (online) for model specifications.

Hypothesis 3. To test the hypothesis that European American and Asian American students will have a higher probability of being identified for gifted services according to national norms than will African American or Latinx students, we examined parameter estimates and their related predicted probabilities for all four models (reading and mathematics with 5% and 15% cutoffs; for details, see Technical Appendix online). We also calculated predicted probabilities for European American students (the reference group), African American students, Asian American students, and Latinx students under national norms, accounting for school-level variables per the previously described approach. We report ORs as the effect sizes for the comparisons of European American students with African American students and European American students with Latinx students.

Hypothesis 4. To test the hypothesis that African American and Latinx students will have a higher probability of being identified for gifted education when building norms are used as compared with national, state, or district norms, we reran the analyses from Hypothesis 3 but changed the reference norms and groups. The specific models are described in the Technical Appendix (online).

Preregistration and Registered Report

With the goals of increasing transparency and confidence in the findings and reducing the overall effort needed to complete the research, we submitted the proposed methods and literature review described here as a Registered Report to *AERA Open* prior to accessing the data. As part of the Registered Report, our proposed study (introduction, literature review, and methods of analysis) was peer reviewed, and reviewer feedback was incorporated into the plan of analysis. This final plan of analysis was then preregistered with the Open Science Framework (https://osf.io/kazy9/) to prevent us from engaging in any questionable research practices (John, Loewenstein, & Prelec, 2012), such as modifying analyses after viewing our data or changing our outcomes to obtain a desired result or to increase the paper's chance of publication. Registered Reports remove desirability bias from the author team as well as the reviewer and editorial teams. By shifting analysis and publication decisions prior to data review, all involved make decisions without being biased by the eventual results. By removing such biases, the

Registered Report process should increase confidence in the internal validity of the study. Moreover, it means that analyses do not have to be conducted multiple times, because reviewers recommend alternative strategies. By shifting the review process prior to data analysis, analyses are only run once. As we note later, any deviation from the preregistered plan of analysis is made clear and justified (e.g., the move to a three-level model instead of the preregistered four-level model in Hypotheses 3 and 4).

Results

Hypothesis 1

Table 2 presents the number of students from each student subgroup identified under each normative criterion level for the 5% and 15% cutoffs in reading and mathematics, as well as enrollment relative to base rate statistics. The results show consistent support for our general hypothesis that more proximal norm groups lead to more racially and ethnically representative populations of identified gifted students. Tables 3 and 4 present change in student representation under each norm cutoff level when compared with national norms for reading and mathematics, respectively. For example, in reading, under building norms for the 5% cut score, African Americans were 238% more represented under building norms than they were under national norms (.76 vs. .22). They remained disproportionately underrepresented under both, but the increase from .22 to .76 far exceeds our a priori criterion for a meaningful improvement. The numbers are relatively similar in mathematics (Table 4). Under the 5% criterion, transitioning from national to building norms led to a 300% increase for African American students (.15 vs. .60), while Latinx students saw a 170% increase in representation (.24 vs. .64).

The effect of more proximal normative criteria on proportionality was less pronounced at the 15% cut score than at the 5%. Although use of district and building norms resulted in substantial increases in proportionality for African American and Latinx students well beyond our 20% a priori criterion, the magnitude of the change in proportion was greater in all cases at the 5% criterion for reading and mathematics. Similarly, although broadening the cut score from 5% to 15% under national norms did increase the number of African American and Latinx students identified as gifted (see Table 2), the change was relatively small when compared with the use of more proximal norm criteria.

Figures 1 and 2 present the percentage identified in reading from each student subgroup under the 5% and 15% criteria, respectively, while Figure 3 shows the percentage changes in representation ratios by ethnicity and cutoff score. Figures 4–6 present the same information for mathematics. Moving from left to right within these figures shows the change in proportion of each group identified under the various normative criteria. Two themes are immediately clear across all these figures. First, national and state norms result in similar proportions of each subgroup being identified, suggesting that the use of state norms would have little to no effect on the size of each group identified. Second, with the exception of the national + building criterion (see Hypothesis 2), more proximal norms led almost every subgroup to become closer to equitable representation.

In summary, the results for reading and mathematics generally support Hypothesis 1: building norms produced an identified gifted population nearer to proportional representation than national norms for African American and Latinx students. All these values exceeded our a priori 20% criterion for meaningful change.

Exploratory results for Hypothesis 1. Although the representation of students from African American and Latinx subgroups increased under more proximal norms, the representation of European American and Asian American students decreased under more proximal norms. For example, under a 5% cutoff, Asian American representation decreased from 2.30 to 1.37 in reading. Similarly, European American student representation decreased from 1.29 to 1.12. Similar decreases were observed at the 15% cutoff as well as both cutoffs in mathematics. Regardless, in all cases, these groups were still disproportionately overrepresented in the identified gifted population.

Hypothesis 2

Whereas Hypothesis 1 primarily assessed the representation change in the shift from national to building norms, Hypothesis 2 assessed changes with a combination of national or building norms with the *or* rule (see Table 2). Results support Hypothesis 2 that using national + building norms would create a gifted-identified population more representative of the larger K–12 population than national norms only but not as representative as building norms. Under the national + building norm, more African American and Latinx students would be identified as gifted than under any other norm criterion. However, disproportionality within the identified population actually becomes *worse* than under building norms because proportionately more European American and Asian American students are identified under the national + building norm (see Figures 3 and 6 as well as the Figures Appendix online). As stated earlier, in almost every case, representation is closest to 1.0 under building norms. The exception is the "other" category at the 5% cut score, where national + building is better than building norms alone in achieving proportionality. The reading and mathematics results are similar, with national + building norms resulting in better proportionality for European American, Asian American, African American, and Latinx students than national norms alone (with building-level norms exceeding both).

TABLE 2

Descriptive Statistics and RIs for the Number and Percentage of Students Identified Under the Different Norms

	Reading						Mathematics					
	5% cutoff			15% cutoff			5% cutoff			15% cutoff		
Students: Norms	*n*	%	RI	*n*	%	RI	*n*	%	RI	*n*	%	RI
African American												
Total	436,596	13.80		436,596	13.80		451,074	13.65		451,074	13.65	
National	2,842	3.10	0.22	22,651	4.97	0.36	2,288	2.04	0.15	15,701	3.62	0.27
State	3,129	3.40	0.25	24,168	5.30	0.38	2,739	2.44	0.18	17,734	4.08	0.3
District	7,493	7.73	0.56	39,143	8.60	0.62	6,515	5.78	0.42	32,737	7.44	0.55
Building	10,138	10.50	0.76	47,849	10.56	0.77	9,097	8.15	0.6	41,841	9.45	0.69
National + building	11,013	8.03	0.58	52,349	8.81	0.64	9,888	6.08	0.45	45,546	7.74	0.57
Asian American												
Total	121,415	3.84		121,415	3.84		126,879	3.84		126,879	3.84	
National	8,082	8.82	2.3	30,494	6.69	1.74	14,641	13.03	3.39	36,633	8.44	2.2
State	7,715	8.37	2.18	29,623	6.49	1.69	14,363	12.77	3.33	36,215	8.33	2.17
District	5,937	6.12	1.59	24,335	5.35	1.39	10,876	9.65	2.51	29,913	6.80	1.77
Building	5,062	5.24	1.37	22,214	4.90	1.28	9,217	8.26	2.15	27,488	6.21	1.62
National + building	9,654	7.04	1.83	34,768	5.85	1.53	16,858	10.36	2.7	41,728	7.09	1.85
European American												
Total	1,700,884	53.77		1,700,884	53.77		1,734,508	52.48		1,734,508	52.48	
National	63,708	69.53	1.29	312,285	68.49	1.27	77,016	68.52	1.31	304,818	70.22	1.34
State	64,948	70.50	1.31	314,037	68.84	1.28	77,307	68.76	1.31	303,971	69.90	1.33
District	63,183	65.15	1.21	289,688	63.64	1.18	73,963	65.63	1.25	283,540	64.47	1.23
Building	58,200	60.28	1.12	272,174	60.08	1.12	69,169	61.95	1.18	270,123	61.00	1.16
National + building	85,489	62.36	1.16	368,202	62.00	1.15	103,132	63.40	1.21	372,355	63.28	1.21
Latinx												
Total	404,941	12.80		404,941	12.80		443,298	13.41		443,298	13.41	
National	3,094	3.38	0.26	22,456	4.92	0.38	3,555	3.16	0.24	20,824	4.80	0.36
State	3,041	3.30	0.26	22,003	4.82	0.38	3,738	3.32	0.25	22,118	5.09	0.38
District	6,168	6.36	0.5	34,625	7.61	0.59	7,145	6.34	0.47	36,861	8.38	0.62
Building	8,374	8.67	0.68	41,588	9.18	0.72	9,551	8.55	0.64	44,740	10.10	0.75
National + building	9,516	6.94	0.54	46,815	7.88	0.62	10,896	6.70	0.5	49,836	8.47	0.63
Other												
Total	499,545	15.79		499,545	15.79		549,292	16.62		549,292	16.62	
National	13,896	15.17	0.96	68,095	14.93	0.95	14,900	13.26	0.8	56,096	12.92	0.78
State	13,293	14.43	0.91	66,368	14.55	0.92	14,289	12.71	0.76	54,837	12.61	0.76
District	14,204	14.65	0.93	67,388	14.80	0.94	14,197	12.60	0.76	56,765	12.91	0.78
Building	14,782	15.31	0.97	69,181	15.27	0.97	14,618	13.09	0.79	58,634	13.24	0.8
National + building	21,423	15.63	0.99	91,737	15.45	0.98	21,883	13.45	0.81	78,994	13.42	0.81
All students												
Total	3,163,381			3,163,381			3,305,051			3,305,051		
National	91,622	2.90		455,981	14.41		112,401	3.40		434,072	13.13	
State	92,126	2.91		456,199	14.42		112,436	3.40		434,876	13.16	
District	96,985	3.07		455,179	14.39		112,696	3.41		439,817	13.31	
Building	96,556	3.05		453,006	14.32		111,651	3.38		442,827	13.40	
National + building	137,095	4.33		593,871	18.77		162,657	4.92		588,459	17.80	

Note. The *n* values represent the total number of students and the number of students identified as gifted. Percentages represent the percentage of that population who is of that race/ethnicity. RI = representation index.

TABLE 3

Change in Representation Rate in Reading by Norm and Race/Ethnicity With National Norms as the Reference Group

	5%					15%				
Norms	AA	Asian Am	Euro Am	Latinx	Other	AA	Asian Am	Euro Am	Latinx	Other
State	9	−5	1	−2	−5	0	−3	1	−2	−3
District	**149**	−31	−6	**88**	−3	**62**	−20	−7	**54**	−1
Building	**238**	−41	−13	**157**	1	**99**	−27	−12	**86**	2
National + building	**159**	−20	−10	**106**	3	**66**	−12	−9	**60**	3

Note. Values are presented as percentages. Bold indicates that the change in representation index exceeded our a priori criterion of ≥20% for a meaningful increase. AA = African American.

TABLE 4

Change in Representation Rate in Mathematics by Norm and Race/Ethnicity With National Norms as the Reference Group

	5%					15%				
Norms	AA	Asian Am	Euro Am	Latinx	Other	AA	Asian Am	Euro Am	Latinx	Other
State	20	−2	0	5	−4	−16	−1	0	6	−2
District	**184**	−26	−4	**100**	−5	**52**	−19	−8	**75**	0
Building	**300**	−37	−10	**170**	−1	**94**	−26	−13	**111**	2
National + building	**199**	−20	−7	**112**	1	**59**	−16	−10	**77**	4

Note. Values are presented as percentages. Bold indicates that the change in representation index exceeded our a priori criterion of ≥20% for a meaningful increase. Also, the change for AA under state norms was 20% after rounding; thus, the percentage is not bold. AA = African American.

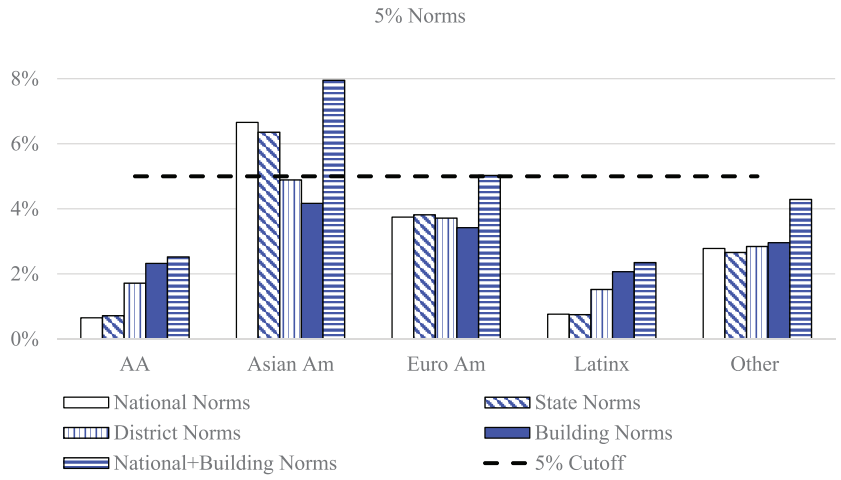

FIGURE 1. *Proportion of each race/ethnicity that was identified as gifted in reading by scope of norm at 5% cutoff. AA = African American.*

Exploratory results for Hypothesis 2. One result that was not part of Hypothesis 2 is that in most cases (e.g., Asian American students at either cut score), district norms turned out to be the second-most proportional norm (after building norms) for achieving proportionality and were similar to national + building norms. This can best be seen in Figure 2A in the Technical Appendix (online).

Hypotheses 3 and 4

For Hypotheses 3 and 4, we shifted from assessing the impact of different norms in the aggregate to the expected impact of using different norms at a typical school. These results can be interpreted as the expected changes in representation based on various norms at a school with an average

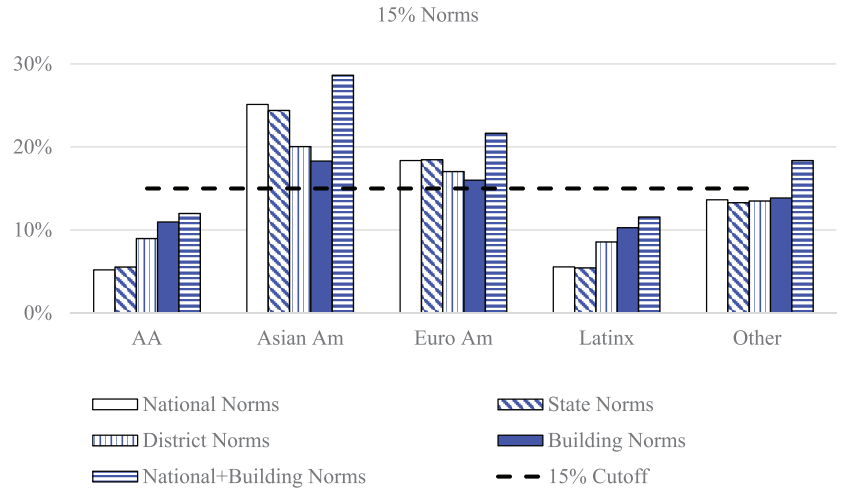

FIGURE 2. *Proportion of each race/ethnicity that was identified as gifted in reading by scope of norm at 15% cutoff. AA = African American.*

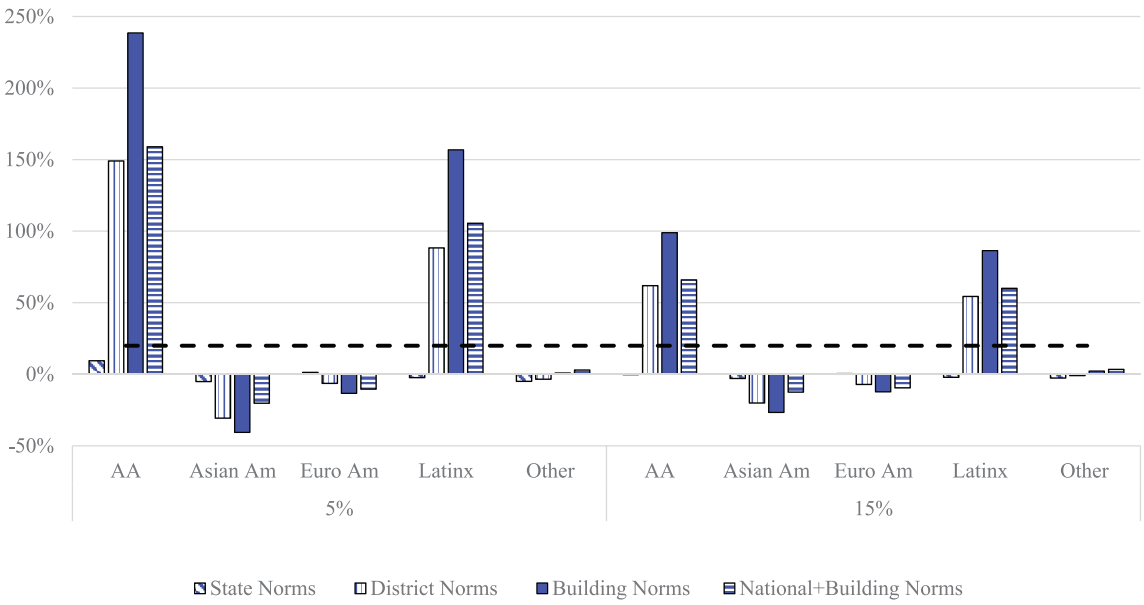

FIGURE 3. *Percentage change in representation indices in reading (national as reference norm) by ethnicity and cutoff. AA = African American.*

number of minority students (30% in our data) after adjustment of the parameter estimates for public/private school status and school locale (i.e., urban, suburban, rural, or town).

We originally planned to test Hypotheses 3 and 4 using four-level models: repeated measures (level 1) within student (level 2) within schools (level 3) within districts (level 4). However, the four-level models would not converge or were completed with errors. Upon inspection, of the 3,424 districts in the reading file and 3,469 districts in the mathematics file, just under 70% (n = 2,367 and = 2,406, respectively) of those districts only had one school, making the district and school levels functionally equivalent. Thus, we removed the fourth level (district) and proceeded with three-level models (repeated measures within student nested within schools). All models then completed without errors.

Hypothesis 3. With this hypothesis, we examined the OR of being identified on the basis of national norms by race/ethnic groups. The odds of being identified for gifted services per national norms for European Americans are provided in the first columns of Table 5, which served as the reference group for the other three groups. Across all four models, Asian American students had a higher probability than European

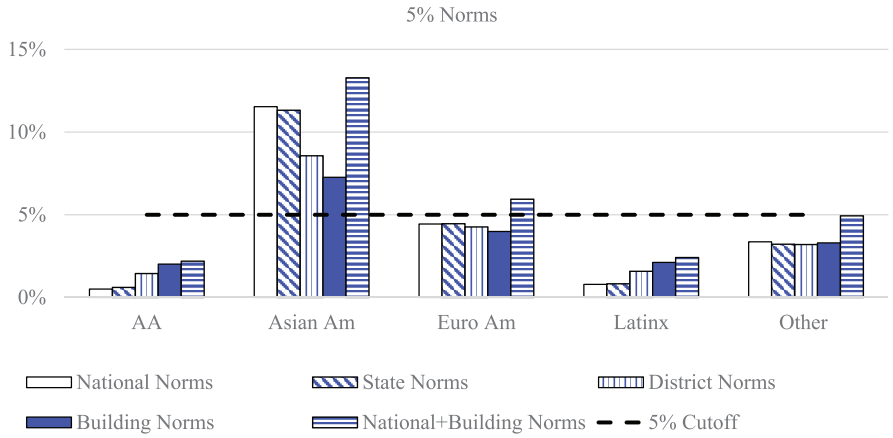

FIGURE 4. *Proportion of each race/ethnicity that was identified as gifted in mathematics by scope of norm at 5% cutoff. AA = African American.*

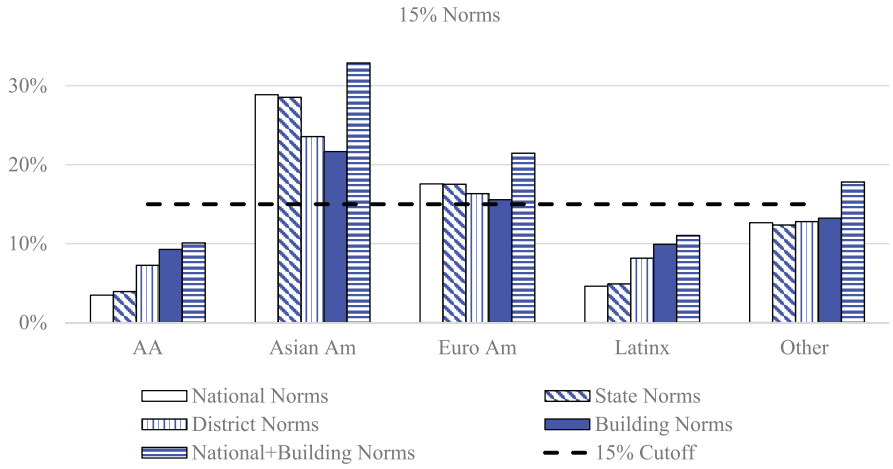

FIGURE 5. *Proportion of each race/ethnicity that was identified as gifted in mathematics by scope of norm at 15% cutoff. AA = African American.*

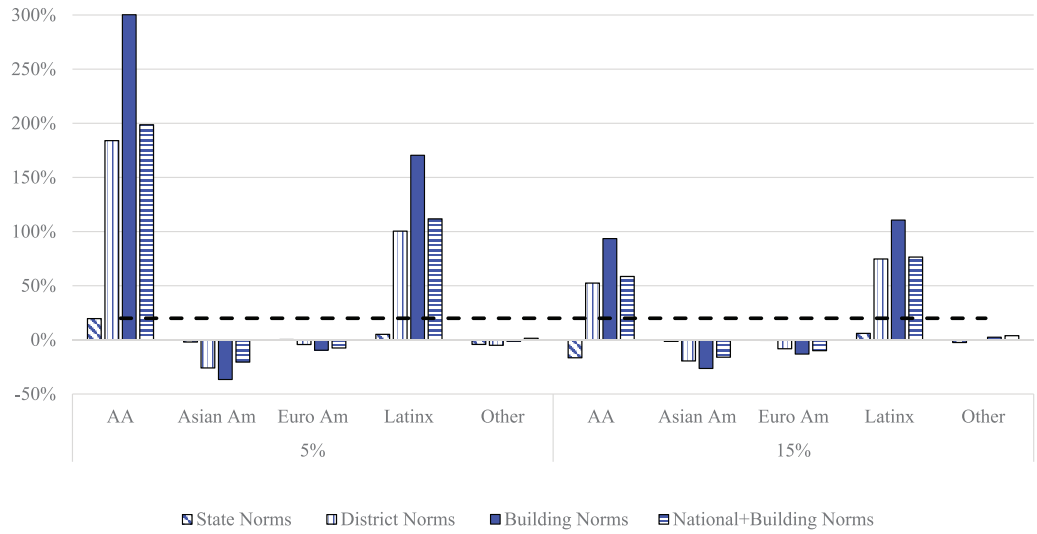

FIGURE 6. *Percentage change in representation indices in mathematics (national norm as reference) by ethnicity and cutoff. AA = African American.*

TABLE 5
Model Estimates and Odds Ratios by Student Subgroup: National Norms

	European American		Asian American		African American		Latinx	
Test: Cutoff	Reference coefficient	OR	ΔCoefficient	ΔOR	ΔCoefficient	ΔOR	ΔCoefficient	ΔOR
Reading								
5%	−4.16	0.02	0.32	1.38	−1.42	0.24[m]	−1.23	0.29[m]
15%	−2.26	0.10	0.28	1.33	−1.42	0.24[l]	−1.3	0.27[m]
Mathematics								
5%	−4.18	0.02	0.75	2.12[s]	−1.63	0.20[l]	−1.23	0.29[m]
15%	−2.44	0.09	0.64	1.90[s]	−1.63	0.20[l]	−1.28	0.28[m]

Note. Model parameter estimates and related ORs of European Americans (reference group) and the changes in the estimates and related changes in the ORs for Asian American, African American, and Latinx students identified under national norms controlling for school level variables. The superscript *s, m*, and *l* denote small, medium, and large Cohen's *d* effect sizes, respectively, per Chen, Cohen, and Chen (2010) and Yuanyuan Lu and Henian Chen (personal communication, November 26, 2018). All coefficients, $p < .001$. OR = odds ratio.

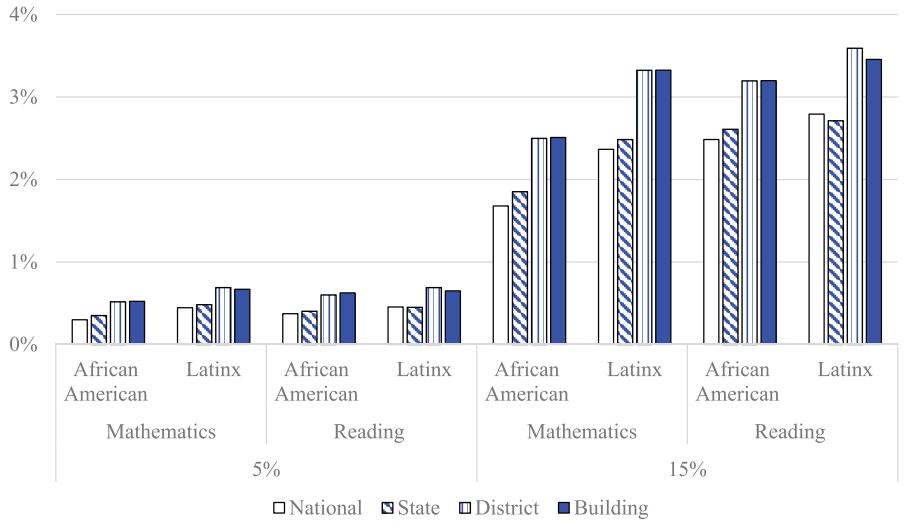

FIGURE 7. *Model-implied probabilities of being identified as gifted according to national, state, district, and building norms (at 5% and 15% cutoffs) for African American and Latinx students.*

American students of being identified as gifted with national norms (i.e., ORs all >1). However, African American and Latinx students showed a lower probability of being identified as gifted than European American and Asian American students for reading and mathematics and at the 5% and 15% cutoff thresholds (i.e., changes in the ORs all <1).

As reported in Table 5, the Cohen's *d* equivalent for the changes in the OR for African American and Latinx students relative to European American students was medium to large (Chen et al., 2010; Yuanyuan Lu and Henian Chen, personal communication, November 26, 2018). Of note, the Cohen's *d* effect size is based on the expected rate of incidence (5% and 15%), so the same OR under different cutoffs may not be considered the same size effect. All final model parameter estimates are reported in the Results Appendix (online).

Figure 7 illustrates the probabilities of being identified for gifted services for schools with an average percentage of African American and Latinx students (30%), controlling for public/private status and urbanity of setting. The patterns across reading and mathematics are the same: African American and Latinx students consistently have lower probabilities of being identified under national and state norms than under district or building norms. This is in line with national data on identification rates (Peters et al., 2018) as well as the results from Hypotheses 1 and 2.

Hypothesis 4. In Hypothesis 4, we evaluated the change in probability of being identified as gifted, similar to Hypothesis 3, but with the reference norm criteria changed to building and the reference subgroup changed to African American students. This was done to directly test the hypothesis that African American and Latinx students would have higher probabilities of identification under building norms. The same set of models was run with Latinx students as the

TABLE 6

Change in Model Coefficients and Odds Ratios of Identification for African American and Latinx Students Under Various Normative Criteria Compared to Building Norms

	Building		National change		State change		District change	
Test: Cutoff	Coefficient	OR	ΔCoefficient	ΔOR	ΔCoefficient	ΔOR	ΔCoefficient	ΔOR
African American								
Reading								
5%	−5.07***	0.006	−0.52***	0.59[s]	−0.44***	0.64[s]	−0.04	0.95
15%	−3.41***	0.030	−0.26***	0.76	−0.21***	0.81	−0.001	0.99
Mathematics								
5%	−5.25***	0.005	−0.56***	0.57[s]	−0.4***	0.67	−0.01	0.99
15%	−3.66***	0.025	−0.41***	0.66[s]	−0.31***	0.73	−0.004	1.00
Latinx								
Reading								
5%	−5.03***	0.006	−0.36***	0.7	−0.37***	0.69	0.06	1.07
15%	−3.33***	0.040	−0.22***	0.8	−0.25***	0.78	0.04	1.03
Mathematics								
5%	−5.00***	0.006	−0.41***	0.66	−0.33***	0.72	0.03	1.03
15%	−3.37***	0.030	−0.35***	0.7	−0.30***	0.74	−0.0004	1.00

Note. The model parameter estimates and related ORs comparing the probability of African American and Latinx students being identified under building norms as compared with national, state, and district norms controlling for school level variables. The superscript *s* denotes a small Cohen's *d* effect size per Chen, Cohen, and Chen (2010) and Yuanyuan Lu and Henian Chen (personal communication, November 26, 2018). OR = odds ratio.
*** $p < .001$.

reference group to assess their change in probability for being identified under each norm as compared with building.

The results reported in Table 6 in the Building column indicate the OR for being identified for gifted services on the basis of building norms at a school with an average percentage of African American and Latinx students after controlling for public/private school status and school locale. For reading and mathematics, the probability of being identified under national and state norms for African American and Latinx students was smaller than under building norms, as evidenced by the changes in the OR for national and state norms all being <1 (see Table 6 and Figure 7). However, although statistically significant, the effect sizes were negligible or small. Additionally, for African American and Latinx students, there was no difference in the OR for being identified according to building norms versus district norms after controlling for school-level variables. All the changes in the OR for using district relative to building norms were essentially nonexistent. Thus, Hypothesis 4 was not fully supported: there were no differences between building and district norms, but national norms and state norms identified fewer African American and Latinx students than building norms. All final model parameter estimates are reported in the Results Appendix (online).

In summary, Hypotheses 1–3 were fully supported by the data, and Hypothesis 4 was partially supported by the data. After controlling for school-level variables, national and state norms did identify fewer African American and Latinx students than building norms, but district norms did not identify fewer African American and Latinx students than building norms. The overall message is clear: the more proximal the norm, the more diverse the students who are identified for gifted services. However, the magnitude of the change will vary across schools.

Discussion

In Hypothesis 1, we predicted at least a 20% improvement in the representation index for African American and Latinx students being identified as gifted according to building norms versus national norms. Our findings supported this hypothesis. This is in line with prior research (Lohman, 2005; Peters & Engerrand, 2016; Plucker & Peters, 2016) suggesting local norms as a way to diversify gifted and talented populations. More broadly, we found that shifting identification criteria from national norms to any more proximal norming group (with the exception of state norms) appeared to lead to a meaningful increase (i.e., >20% gain) in gifted representation rates for African American and Latinx students across mathematics and reading.

A consequential implication (not part of our a priori hypotheses or predictions) of any shift to nonnational normative criteria would be a nontrivial decrease in representation of Asian American students in gifted programs. For example, by using a 5% cutoff and building norms, Asian American student representation in gifted programs would

decline 41% in reading and 37% in mathematics. However, as shown in Table 2, these student groups would still be disproportionately overrepresented in gifted programs relative to their proportions of the student population (1.37 and 2.15 times as likely to be identified as gifted under building norms in reading and mathematics, respectively). European American students appear to show similar declines, but as shown in Table 2, European American students would remain disproportionately overrepresented in gifted programs relative to their proportion among the overall student population. This decline would fall below our a priori 20% threshold for meaningful change.

In support of Hypothesis 2, we found that using national + building norms yielded proportionality closer to parity (i.e., 1.0) for European American, Asian American, African American, and Latinx students than national norms did, and this held for reading and mathematics. Additionally, building-level norms exceeded national norms alone and national + building norms in terms of proportionality. The national + building option can be seen as a compromise between the extremes of national and building norms. Under national + building norms, fewer Asian American or European American students would be seen as "losing" eligibility because they would remain identified under the national norm pathway, while additional African American and Latinx students would be identified. This compromise comes at the cost of identifying the largest number of students. Table 2 shows that under the 5% cutoff for reading, 91,622 and 96,556 students were identified via the national and building norm criteria, respectively, whereas 137,095 students were identified under the national + building criterion—a 40% increase over building norms. Although the expense of using different norm criteria itself would not be high, we imagine that a ≥40% increase in the population eligible to receive gifted services would be significant for any school district in terms of additional resources needing to be allocated to this area. Another alternative would be to phase in the nonnational normative criteria over time—for example, one grade level at a time. This way, students identified with national norms would age out of the system, with incoming students being identified with more-local normative criteria.

School-level factors influence the degree to which a move from national to building norms increases diversity within gifted education. A move to building norms in schools with a typical proportion of minority students (about 30%) does not have as large an effect as it would in a building with a larger proportion of minority students. This points to an important implication, which is that building norms will not increase the diversity within every building's gifted population. Rather, building norms have the greatest effect (a) on the aggregate population diversity and (b) in schools with larger-than-average populations of minority students. Thus, although our data suggest that implementation of building norms would yield a massive increase in the number of African American and Latinx students in advanced learning programs across the United States, policy makers should not presume that building norms will have such effects in every school. Implementing local norms is not a panacea for addressing all systemic causes of underrepresentation in gifted education.

Two important caveats to our analyses are (a) the need for universal consideration and (b) a caution to educators that use of local norms need not automatically result in loss of services for some students previously identified with the use of less proximal norms. Regarding universal consideration, districts will find it impossible to take the top 5% or 15% of any group if <100% of the group is tested. This is an important caveat simply because universal consideration of an entire grade for gifted programming eligibility is still relatively rare in U.S. schools, due in part to little state financial support for such systems (Plucker, Glynn, Healey, & Dettmer, 2018). This is not a limitation of the current study's analyses but instead represents a challenge to their broader implementation.

Regarding loss of services that could be seen in a transition from one norm criterion to another (e.g., national to building), the few districts in the country that are beginning to share experiences with local norms implementation generally report resistance from parents whose children (presumably European American, Asian American, and/or upper income) lose services as a result. District leadership often appears surprised by this political blowback, but it is to be expected whenever students lose services. An approach that expanded the number of students receiving advanced learning services (i.e., through a combination of national/state/district and local norms) would require more resources but likely result in less controversy. This is an important consideration before a district moves forward with any new normative criteria.

Conclusion

Disproportionality rates along racial/ethnic lines have been reported in numerous educational areas, including gifted identification. Such disproportionality is particularly problematic if/when the students who remain unidentified would benefit from placement into gifted education programming. The current results, in tandem with previous findings, suggest that transitioning from national/state reference norms to local building norms for gifted identification would substantially reduce group differences in rates of gifted identification. Practically, such a shift would also help schools identify a more consistent number of students. This approach also would constitute a shift toward difference from within-building peers as the justification for gifted services. The metric for success of the gifted identification process would be increased learning outcomes that

arise from placing students into environments that can better meet their specific learning needs. Such a change in identification policy would likely require schools to provide additional teacher training and possibly reallocate space and staff, but any such changes to serve a larger and more representative gifted student population should be seen as necessary expenditures in the service of improved educational equity.

The results in this article can be used by school leaders when contemplating the adoption of different norming strategies in their gifted education identification systems. First, in districts with little residential segregation (de facto or otherwise) and similar demographics across their schools, using any level of norms will likely produce a similar pool of identified students. Second, in districts with considerable residential segregation and dissimilar school demographics, using school-based norms will identify more African American and Latinx students (and, although not directly addressed in this study, almost certainly more students from low-income backgrounds). Third, if a district does not expand programming and holds the number of identified students constant, moving from national to local norms will result in some previously identified students losing services, and a negative parent and student reaction should be expected. An alternative is to expand the number of students receiving services by using both school-based and national/state/district norms, which will not improve disproportionality as dramatically as local norms alone but should sharply increase the number of previously underserved African American and Latinx students eligible for gifted services.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by research funds provided to the fifth author by Johns Hopkins University and through the Richard and Veronica Telfer Endowed Faculty Fellowship to the first author.

Notes

1. Although race is not a construct with clearly defined boundaries, we believe that race and the underlying power dynamics of racial categories matter. We purposefully vary racial terms (e.g., African American and European American, rather than African American/Black and White). We recognize that umbrella terms such as Native American incorporate numerous subgroups, as do the descriptors Asian, Hispanic, and Latinx. We consciously have changed or rearranged some terms used in national databases, as these may demonstrate a bias hierarchy. We also follow American Psychological Association guidelines and arrange lists of racial/ethnic categories or results either alphabetically or numerically.

2. All data are from the NAEP Data Explorer, https://nces.ed.gov/nationsreportcard/naepdata/dataset.aspx. Data run May 22, 2018.

ORCID iDs

S. J. Peters https://orcid.org/0000-0003-2459-3384

K. Rambo-Hernandez https://orcid.org/0000-0001-8107-2898

M. C. Makel https://orcid.org/0000-0002-3837-0088

M. S. Matthews https://orcid.org/0000-0003-1695-2498

References

American Educational Research Association, American Psychological Association, & National Council for Measurement in Education. (2014). *The standards for educational and psychological testing*. Washington, DC: Author.

Arizona State Legislature. (n.d.). *Gifted pupils; scope and sequence; annual funding*. Retrieved from https://www.azleg.gov/ars/15/00779-02.htm

Benjamin, D. J., Berger, J. O., Johannesson, M., Nosek, B. A., Wagenmakers, E. J., Berk, R., . . . Cesarini, D. (2018). Redefine statistical significance. *Nature Human Behaviour*, *2*, 6–10. doi:10.1038/s41562-017-0189-z

Callahan, C. M., Moon, T. R., & Oh, S. (2014). *National survey of gifted programs: Executive summary*. Retrieved from http://www.nagc.org/sites/default/files/key%20reports/2014%20Survey%20of%20GT%20programs%20Exec%20Summ.pdf

Carman, C. A., Walther, C. A. P., & Bartsch, R. A. (2018). Using the Cognitive Abilities Test (CogAT) 7 Nonverbal Battery to identify the gifted/talented: An investigation of demographic effects and norming plans. *Gifted Child Quarterly*, *62*, 193–209. doi:10.1177/0016986217752097

Chen, H., Cohen, P., & Chen, S. (2010) How big is a big odds ratio? Interpreting the magnitudes of odds ratios in epidemiological studies. *Communications in Statistics—Simulation and Computation*, *39*, 860–864, doi:10.1080/03610911003650383

Colorado Department of Education. (2016). *Gifted identification*. Retrieved from https://www.cde.state.co.us/gt/idguidebook

Cruz, R. A., & Rodl, J. E. (2018). An integrative synthesis of literature on disproportionality in special education. *The Journal of Special Education*, *52*, 50–63. doi:10.1177/0022466918758707

Davies, H. T. O., Crombie, I. K., & Tavakoli, M. (1998). When can odds ratios mislead? *BMJ*, *316*(7136), 989–991. doi:10.1136/bmj.316.7136.989

Garland, S. (2013, March 14). Ending racial inequality in gifted education. *The Hechinger Report*. Retrieved from http://hechingerreport.org/ending-racial-inequality-in-gifted-education/

Georgia Department of Education. (2017). *Georgia resource manual for gifted education services: 2017–2018*. Retrieved from https://www.gadoe.org/Curriculum-Instruction-and-Assessment/Curriculum-and-Instruction/Documents/Gifted%20Education/Georgia-Gifted-Resource-Manual.pdf

Gregory, A., Cornell, D., & Fan, X. (2011). The relationship of school structure and support to suspension rates for Black and White high school students. *American Educational Research Journal*, *48*, 904–934. doi:10.3102/0002831211398531

Grissom, J. A., & Redding, C. (2016). Discretion and disproportionality: Explaining the underrepresentation of high-achieving student of color in gifted programs. *AERA Open*, *2*(1), 1–25. doi:10.1177/2332858415622175

Gupta-Kagan, J. (2017). The school-to-prison pipeline's legal architecture: Lessons from the Spring Valley incident and its aftermath. *Fordham Urban Law Journal*, *45*, 83–147.

Hamilton, R., McCoach, D. B., Tutwiler, M. S., Siegle, D., Gubbins, E. J., Callahan, C. M., . . . Mun, R. U. (2018). Disentangling the roles of institutional and individual poverty in the identification of gifted students. *Gifted Child Quarterly*, *62*, 6–24. doi:10.1177/0016986217738053

John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the provenance of questionable research practices with incentives for truth telling. *Psychological Science*, *23*, 524–532. doi:10.1177/0956797611430953

Kingsbury, G. G., & Wise, S. L. (2011). Creating a K–12 adaptive test: Examining the stability of item parameter estimates and measurement scales. *Journal of Applied Testing Technology*, *12*, 3–5.

Lakens, D. (2014). Performing high-powered studies efficiently with sequential analyses: Sequential analyses. *European Journal of Social Psychology*, *44*, 701–710. doi:10.1002/ejsp.2023

Lakin, J. M. (2018). Making the cut in gifted selection: Score combination rules and their impact on program diversity. *Gifted Child Quarterly*, *62*, 210–219. doi:10.1177/0016986217752099

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. doi:10.4219/jeg-2005-341

Lohman, D. F. (2009). Identifying academically talented students: Some general principles, two specific procedures. In L. V. Shavinina (Ed.), *International handbook on giftedness* (pp. 971–997). Berlin, Germany: Springer Science + Business Media.

McBee, M. T., Peters, S. J., & Miller, E. M. (2016). The impact of the nomination stage on gifted program identification: A comprehensive psychometric analysis. *Gifted Child Quarterly*, *60*, 258–278. doi:10.1177/0016986216656256

McBee, M. T., Peters, S. J., & Waterman, C. (2014). Combining scores in multiple-criteria assessment systems: The impact of the combination rule. *Gifted Child Quarterly*, *58*, 69–89. doi:10.1177/0016986213513794

McCall, M., Kingsbury, G., & Olson, A. (2004). *Individual growth and school success: A technical report from the NWEA growth research database*. Portland, OR: Northwest Evaluation Association.

Morgan, P. L., Farkas, G., Hillemeier, M. M., & Maczuga, S. (2017). Replicated evidence of racial and ethnic disparities in disability identification in U.S. schools. *Educational Researcher*, *46*, 305–322. doi:10.3102/0013189X17726282

National Association for Gifted Children. (2015). *State of the states in gifted education: 2014–2015*. Washington, DC: Author.

Northwest Evaluation Association. (2011). *MAP V3 for Colorado—Instructional statements*. Retrieved from http://www.nwea.org

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

Peters, S. J., & Gentry, M. (2012). Group-specific norms and teacher rating scales: Implications for underrepresentation. *Journal of Advanced Academics*, *23*, 125–144. doi:10.1177/1932202X12438717

Peters, S. J., Gentry, M., Whiting, G. W., & McBee, M. T. (2019). Who gets served in gifted education? Demographic proportionality and a call for action. *Gifted Child Quarterly*. Advance online publication. doi:10.1177/0016986219833738

Peters, S. J., Rambo-Hernandez, K., Makel, M. C., Matthews, M. S., & Plucker, J. A. (2017). Should millions of students take a gap year? Large numbers of students start the school year above grade level. *Gifted Child Quarterly*, *61*, 229–238. doi:10.1177/0016986217701834

Plucker, J., Glynn, J., Healey, G., & Dettmer, A. (2018). *Equal talents, unequal opportunities: A report card on state support for academically talented low-income students* (2nd ed.). Retrieved from https://www.jkcf.org/wp-content/uploads/2018/06/FINAL_2018_JKCF_-_Equal_Talents_Unequal_Opportunities_-_Web_version.pdf

Plucker, J. A., Hardesty, J., & Burroughs, N. (2013). *Talent on the sidelines: Excellence gaps and America's persistent talent underclass*. Storrs, CT: University of Connecticut, Center for Education Policy Analysis.

Plucker, J. A., & Peters, S. J. (2016). *Excellence gaps in education: Expanding opportunities for talented students*. Cambridge, MA: Harvard Education Press.

Plucker, J. A., & Peters, S. J. (2018). Closing poverty-based excellence gaps: Conceptual, measurement, and educational issues. *Gifted Child Quarterly*, *62*, 56–67. doi:10.1177/0016986217738566

Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., Congdon, R. T., & du Toit, M. (2016). *HLM7: Hierarchical linear and nonlinear modeling*. Skokie, IL: Scientific Software International.

Renzulli, J. S. (1978). What makes giftedness? Reexamining a definition. *Phi Delta Kappan*, *60*, 180–184.

Renzulli, J. S. (2005). *Equity, excellence, and economy in a system for identifying students in gifted education programs: A guidebook* (Publication No. RM05208). Storrs: University of Connecticut, The National Research Center on the Gifted and Talented.

Siegle, D., McCoach, D. B., Gubbins, E. J., Long, D., & Hamilton, R. (2018, January). *Gifted identification gap: When just as good is not good enough*. Report presented at the Institute of Educational Sciences Annual Principal Investigators Meeting, Arlington, VA.

Stanley, J. C. (2000). Helping students learn only what they don't already know. *Psychology, Public Policy, and Law*, *6*, 216–222.

State of Tennessee. (2017). *Standards for special education evaluation and eligibility*. Retrieved from https://www.tn.gov/content/dam/tn/education/special-education/eligibility/se_eligibility_intellect_gift.pdf

U.S. Department of Education. (1993). *National excellence: The case for developing America's talent*. Washington, DC: U.S. Government Printing Office.

U.S. Department of Education. (2016). *Status and trends in the education of racial and ethnic groups 2016*. Retrieved from https://nces.ed.gov/pubs2016/2016007.pdf

Wang, S., McCall, M., Jiao, H., & Harris, G. (2012, April). *Construct validity and measurement invariance of computerized adaptive testing: Applications to Measures of Academic Progress (MAP) using confirmatory factor analysis*. Paper presented at the meeting of the American Educational Research Association, Vancouver, Canada.

Williams, R. B., Bryant-Mallory, D., Coleman, K., Gotel, D., & Hall, C. (2017). An evidence-based approach to reducing disproportionality in special education and discipline referrals. *Children & Schools*, *39*, 248–251. doi:10.1093/cs/cdx020

Worrell, F. C. (2018). Identifying gifted learners: Utilizing nonverbal assessment. In C. M. Callahan, & H. L. Hertberg-Davis (Eds.), *Fundamentals of gifted education: Considering multiple perspectives* (pp. 125–134). New York, NY: Routledge.

Yaluma, C. B., & Tyner, A. (2018). *Is there a gifted gap? Gifted education in high-poverty schools*. Retrieved from https://edexcellence.net/publications/is-there-a-gifted-gap

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. doi:10.1177/0016986208330564

Authors

SCOTT J. PETERS is an associate professor of educational foundations and the Richard and Veronica Telfer Endowed Faculty Fellow of Education at the University of Wisconsin–Whitewater. His research work focuses on educational assessment, gifted and talented student identification, disproportionality, and educational policy.

KAREN RAMBO-HERNANDEZ is an associate professor of educational psychology in the Department of Learning Sciences and Human Development at West Virginia University. Her research interests include novel applications of multilevel modeling and growth modeling, the assessment of educational interventions to improve STEM education, and access for all students—particularly high-achieving and underrepresented students—to high-quality education.

MATTHEW C. MAKEL is the director of research and evaluation for Duke University's Talent Identification Program. His research focuses on academic talent development and research methods.

MICHAEL S. MATTHEWS is professor and director of the academically and intellectually gifted graduate programs at the University of North Carolina at Charlotte and coeditor of *Gifted Child Quarterly*. His research interests focus on assessment and identification of gifted children, education policy, parenting (including homeschooling) of gifted learners, and gifted and academically advanced learners from diverse backgrounds, particularly English learners.

JONATHAN A. PLUCKER is the Julian C. Stanley Professor of Talent Development at Johns Hopkins University. His work focuses on creativity and intelligence, education policy, and talent development.