UNIVERSITY OF ARKANSAS | College of Education & Health Professions
Education Reform

WORKING PAPER SERIES

Widening the Net: How CogAT and ACT Aspire Compare in Gifted Identification

Al Mansor Helal, Jonathan Wai, Andy Parra-Martinez, Sarah McKenzie, and Dustin Seaton

March 2025

EDRE Working Paper 2025-03

The University of Arkansas, Department of Education Reform (EDRE) working paper series is intended to widely disseminate and make easily accessible the results of EDRE faculty and students' latest findings. The Working Papers in this series have not undergone peer review or been edited by the University of Arkansas. The working papers are widely available, to encourage discussion and input from the research community before publication in a formal, peer reviewed journal. Unless otherwise indicated, working papers can be cited without permission of the author so long as the source is clearly referred to as an EDRE working paper.

Universal Screening in Practice: A Comparative Study of the CogAT and ACT Aspire for Gifted Identification

Al Mansor Helal*
University of Arkansas

Jonathan Wai
University of Arkansas

Andy Parra-Martinez
Mississippi State University

Sarah McKenzie
University of Arkansas & Office of Education Policy (OEP)

Dustin Seaton
Rogers Public Schools

***Corresponding Author:** Al Mansor Helal, University of Arkansas. Email: ah176@uark.edu

Acknowledgments: We are grateful for our collaboration with the Office of Education Policy at the Department of Education Reform at the University of Arkansas, who made access to data possible.

Abstract

Previous research indicates that many academically accomplished students from disadvantaged backgrounds are not identified for gifted and talented (G/T) programs. This study examines a large sample of students (N = 10,508), many of whom took both the Cognitive Abilities Test (CogAT) and the ACT Aspire test. We examined test similarities and differences with an eye to widen the net for G/T identification in practice. This study demonstrates that the ACT Aspire and CogAT have a significant correlation in our sample of r = .59. However, the correlation varies across cohorts noticeably from r = .72 to r = .49. This variation in correlations and inconsistency in the predictive nature of diversity of both tests across cohorts suggest greater caution in the interchangeable use of ACT Aspire and CogAT as indicators. Instead, this suggests the use of more than one test as part of the package for identification.

Keywords: CogAT, ACT Aspire, achievement tests, gifted identification, education policy

Introduction

There is a growing body of literature showing that there is a noticeable inequity in the identification of students for gifted and talented (G/T) services across the U.S. (e.g., Hoxby & Avery, 2013; Plucker & Peters, 2016; Wai & Worrell, 2020). A recent study in Arkansas showed that around 30% of the students who were among the top 5% of scorers on achievement tests were not identified for G/T programs (Tran et al., 2021). Across different states in the U.S., there are various policies to identify G/T students with varying effectiveness in each method. Prior work has indicated that using universal screening (Card & Giuliano, 2016) by leveraging a state standardized achievement test would improve the identification of high achieving talented students, many from low-income backgrounds (Tran et al., 2021). When coupled with "local norms" for selecting the highest achieving students relative to their opportunity to learn (e.g., comparing students at a similar level of family income or local context), this might lead to a wider and more inclusive group of talented students (Peters et al., 2021).

Moreover, as all achievement and reasoning tests that measure verbal and mathematical abilities tend to be highly correlated (e.g., Gomez-Veiga et al., 2018; Naglieri & Ford, 2003; Peng & Kievit, 2020; Wai et al., 2018), schools can use either achievement or reasoning test scores as an objective indicator for G/T identification. And, in practice, schools often need to piggyback on whatever universal data they already have available. However, Lohman (2005) presciently warned about the potential loss of a significant number of academically talented students in G/T programs if schools only consider achievement test scores or ability test scores as identification tool in isolation. Moreover, if schools intend to widen the pool of talented learners from diverse backgrounds, they may use scores of more than one test as an objective indicator ensuring a more universal screening process (Lakin, 2019). Also, because universally provided ability tests are expensive and states already use achievement tests, the use of achievement test scores for initial screening of G/T placement makes sense but also needs more investigation, given this is a lower cost option for universal screening to happen more frequently in practice.

In this study, we examined the overall correlation between the ACT Aspire and CogAT tests in a sample of roughly 10,000 students in Arkansas, in order to understand the suitability of the interchangeable use of these assessments. Further, we investigated how each test varies in identification by students' gender, race/ethnicity, free and reduced lunch (FRL), and English Language Learner (ELL) status. This investigation also explored the results of G/T identification under both assessments on different patterns of diversity. The overarching goal of this study is to explore if the tests differ in predicting diversity and inclusiveness, and what happens if they are used together.

Literature Review

Underrepresentation of Students in G/T Programs

G/T programs are advanced educational opportunities offered for students with high cognitive abilities and potential for high performance (Assouline et al., 2015; Lubinski & Benbow, 2000; Peters et al., 2022), with the broad aim to ensure that students are appropriately developmentally placed where their current learning rate and academic preparation is matched to the rigor and pace of coursework. G/T programs typically offer identified G/T learners accelerated and/or academic enrichment coursework (Subotnik et al., 2011). According to 2017-18 data from the Office of Civil Rights, 6.54% of the total U.S. student population are placed in G/T programs. Researchers focus on how gifted education serves to address the learning needs of G/T students so that they attain optimal educational outcomes (Assouline et al., 2015; Gentry, 2009; Lubinski & Benbow, 2006). This education may happen in many forms: acceleration, curriculum compacting, or enrichment programs in areas of interests (Assouline et al., 2015; Gentry, 2009; Reis & Renzulli, 1991; Wai et al., 2010). To achieve these gifted education objectives, researchers have consistently emphasized the importance of a transparent, research-based, and purposeful identification process (Hodges et al., 2018).

However, due to typically suboptimal identification methods in practice, fairness and equity of identification for G/T programs are often questioned. Yoon and Gentry (2009) expressed concerns about inequitable representation in G/T programs. Card and Giuliano (2016) as well as many other scholars have noted that in the U.S., students from low-income families and historically marginalized groups are significantly underrepresented in gifted education programs (Gentry et al., 2019). Recent studies demonstrate that talented children from low-income families and traditionally underrepresented communities are less likely to reach their full potential when compared to their peers with similar talents, higher income, and other racial/ethnic backgrounds (e.g., Hoxby & Avery, 2013; Plucker & Peters, 2016; Wai & Worrell, 2020). Grissom and Redding (2016), applying a conditional probability approach, found that even among students with high standardized test scores, Black students were less likely to be assigned to gifted services. These authors documented that approximately 7% of White students and 14% of Asian Americans students were identified as gifted by third grade compared to only 2% of Black and 5% of Hispanic students. According to 2017-18 school year enrollment data, gifted programs in public schools throughout the U.S. consisted of 58.4% White students, 18.3% Hispanic students, 9.9% of Asian students and 8.2% of Black students, whereas White students represent 47.3%, Hispanic students 27.2%, Asian students 5.2% and African American students

15.1% of total students enrolled in public schools across the country (Office of Civil Rights, 2018).

Identification Policies

Nomination Versus Universal Screening

Researchers have a wide range of ideas (Giessman et al., 2013; Rasheed, 2020) on how, what, and when to identify gifted learners (Callahan, 2005), though developed reasoning constructs remain central to identifying academic giftedness (e.g., Lohman, 2005; Subotnik et al., 2011; Wai et al., 2018). Acar et al. (2016) categorized gifted identification methods into two broad forms: performance and nonperformance methods. Performance methods consider student test scores, whereas nonperformance methods include all other G/T identification methods not involving any performance-based assessment (Acar et al., 2016). Teacher and parent nominations are one of several examples of nonperformance methods. According to McBee (2006), nomination or referral is a process by which parents or teachers recommend students for screening (or testing) for gifted services. Screening refers to the use of a formal assessment tool for making placement decisions.

However, McBee (2010) argued that nomination is one of the core reasons responsible for such underrepresentation from students from disadvantaged backgrounds (also supported by findings from Grissom and Redding, 2016). Asian and White students were much more likely to be nominated than Black or Hispanic students (McBee, 2010). Students who qualified for FRL were much less likely to be nominated than non FRL students, suggesting that that inequalities in nomination, rather than assessment, may be a source of the underrepresentation of minority and low-SES students in gifted programs. Additionally, because teachers in most schools are of

White middle-class backgrounds, McBee (2010) argued, they may not always notice the signs of giftedness expressed in students from different cultural origins. Thus, part of the underrepresentation issue is exacerbated by potentially suboptimal nomination procedures.

One of the issues embedded in G/T identification is the choice between referral or nomination as the first step to identification, versus the use of universal screening where everyone is tested as the first step (Lakin, 2016). If referral is the first step of identification, then screening occurs only for students that are nominated or referred by their teachers or parents (or other sources). Conversely, if all students in an eligible grade level are administered at least one formal assessment—such as an ability or achievement test—as the first step of identification, this allows all students to have an equal (or closer to equal) chance of being identified based on the same assessments and selection criteria (Lakin, 2016). Card and Giuliano (2016) revealed that testing all students, instead of relying first on referrals, led to a significant increase in the representation of historically underserved students in gifted programs in the sample they studied. Thus, universal screening is one important initial step to address the systemic underrepresentation of students from disadvantaged backgrounds and can enhance diversity in gifted programs (Callahan et al., 2013; Lakin & Lohman, 2011; Wai & Lakin, 2020).

Local Norms

Local norms refer to the identification of students for gifted programs based on norm-referenced interpretations, aligning the selection criteria with the desired level of service (Peters et al., 2021). In contrast to national comparisons or norms produced by test developers, local normative criteria tend to compare students to their immediate peers, e.g., similarity in age, experience, background, learning environment, and nature of the intervention to be provided (Peters et al., 2021: Warne & Larsen, 2022) While tests are normed and often local students do not meet the thresholds of national or specific norms, using local norms can make it easier for teachers to identify the most advanced students with a high likelihood of requiring additional intervention to be challenged, compared to their grade-level peers in the same school (Peters et al., 2021).

Coupled with universal screening, recent research supports the use of local norms to identify gifted learners and ensure greater representation of disadvantaged students in gifted programs (Peters et al., 2019). Peters et al. (2021) showed an increase of 213% of African American students and 213% increase of Latinx students in the gifted programs of a district that adhered to local norms instead of national norms. In another study of 10 U.S. states, Peters et al. (2019) showed a significant improvement of Hispanic and Black students' representation in gifted programs using local norms. Carman et al. (2018) found that Hispanic and Black students qualified for gifted programs at the highest rate when they were tested based on school-level local norms, and at the lowest rate when national norms were used.

Reasoning Versus Achievement Tests

Considering tests, according to Hodges et al. (2018), there are two types of identification methods: traditional and non-traditional. While traditional methods include IQ and standardized achievement tests, nonverbal tests and student portfolios are some examples of nontraditional identification methods. Usually, the identification of G/T students is made through traditional methods, where ability tests are very common. One of the widely used ability tests is the CogAT (Carman, 2018), a measure of verbal, quantitative, and nonverbal reasoning among K-12 students (Lohman & Lakin, 2010). Despite their use in identification, challenges remain with the broad implementation of ability testing (i.e., such tests can be expensive, time consuming, require administration by an expert psychologist, etc.). Additionally, in some cases (depending upon the constructs measured by the test) these assessments are predominantly often suitable for students who have developed verbal skills. This challenge has been partially addressed by the use of non-verbal tests as Lohman (2005) argues that non-verbal ability tests tend to examine more innate talents than learned symbol systems, and test figural reasoning instead of spatial reasoning. Harradine et al. (2014) regarded nontraditional assessments, (i.e., nonverbal reasoning tests) as a reason behind the disproportionate underrepresentation of students of various subgroups in gifted programs. In addition, nonverbal tests, as their strongest limitation, are likely to exclude many academically accomplished students with strong verbal and quantitative reasoning skills and include many students who are not quite ready for the particular type of educational opportunity in schools, where mathematical and verbal skills are necessary (Lohman, 2005).

Another example of traditional methods of G/T identification is the use of achievement tests. Achievement tests are commonly used to measure grade level proficiency in major subjects such as math or science (Sussman & Wilson, 2019). According to May et al. (2009), standardized achievement tests may be suitable for evaluating the impacts of interventions where the goal is to increase grade level proficiency. For example, some use the ACT Aspire as an achievement test, which is a summative assessment of student achievement in English, reading, writing, math, and science (Williamson, 2019). According to the ACT, the ACT Aspire includes a vertically scaled battery of achievement tests designed to measure student growth in a longitudinal assessment system for third through tenth grades (ACT Aspire, n.d.).

All achievement and reasoning tests that measure verbal and mathematical skills tend to be highly correlated (Naglieri et al., 2003; Peng & Kievit, 2020; Wai et al., 2018; Zins & Barnett, 1983). Lohman (2005) demonstrated a correlation of $r = .6$ between a nonverbal ability test and a concurrently administered math achievement test. Naglieri and Ronning (2000) found a correlation of r = .56 between the Naglieri Nonverbal Ability Test (NNAT) and measures of reading in the Stanford Achievement Test Ninth Edition (SAT-9). However, there is also likely variation at the local level, and this variation is important to better understand to help G/T coordinators make more effective G/T placement decisions on the ground. Therefore, in this study we investigated test similarities and differences as well as their prediction of diversity and inclusiveness with an eye to widen the net of G/T identification. We seek to answer the following research questions using a sample of roughly 10,000 students in the state of Arkansas:

Research Question 1 (RQ 1): What is the correlation between the ACT Aspire and CogAT? How does this relationship vary by student demographics: gender, race/ethnicity, FRL, ELL, and G/T statuses?

Research Question 2 (RQ 2): How does diversity in G/T identification differ when using the ACT Aspire test versus the CogAT test?

Methods

Data and Sample:

The data we used were anonymized student-level assessment and demographic data from the Arkansas Department of Education. Publicly available district-level characteristics were then matched with student-level data. Including data from 15 school districts, we examined two cohorts of students assessed during the years 2018 through 2022. Cohort 1 (2018-19 school year) students took the ACT Aspire in 3rd grade and the CogAT in 4th grade in the 2019-20 school year. Students of Cohort 2 took the ACT Aspire and CogAT in the 2021-22 school year during 4^{th} grade.

We used a sample of 10,508 students in total— 5,279 students in Cohort 1 and 5,229 students in Cohort 2. Table 1 reports summary statistics of their demographic characteristics. Across our samples, 53.48% of students were FRL eligible, 49.28% were female, 23.58% received ELL services, 54.62% were white, 3.2% were Black, 34.63% were Hispanic, and 11.33% were identified as G/T.

Table 1.
Student Demographic Characteristics

Variable	Overall		Cohort 1		Cohort 2	
	n	%	*n*	%	*n*	%
Gender						
Female	5,178	49.28	2,605	49.35	2,573	49.21
Male	5,330	50.72	2,674	50.65	2,656	50.79
Race and Ethnicity						
Asian	239	2.28	100	1.9	139	2.66
Black	336	3.2	131	2.38	205	3.92
Hispanic	3,637	34.63	1,895	35.94	1,742	33.31
White	5,408	54.62	2,645	50.17	2,763	52.84
Other race	881	8.39	501	9.49	380	7.27
Educational Characteristics						
ELL	2,477	23.58	1,385	26.25	1,092	20.88
FRL	4,814	53.48	2,039	54.06	2,775	58.09
G/T	1,191	11.33	544	10.31	647	12.37

Note. Data for FRL status in Cohort 1 was not provided by one school district due to privacy concerns.

Empirical Approach

We used Pearson correlations to assess the associations between the ACT Aspire and CogAT tests. We employed Ordinary Least Squares (OLS) regressions to examine whether the correlations between tests remained invariant when controlling for demographics, i.e., students' gender, race, FRL, ELL, student academic status, and G/T status. We utilized the following model:

$$Y_{it} = \beta_0 + \beta_1 CogAT_{it} + X_i + \mu_{tg} + e_i$$

The subscript *i* represents each student who took both the ACT Aspire and CogAT tests in the data from 2018 through 2022. Y_i denotes the standardized ACT Aspire scores based on English Language Arts (ELA) and math scale scores, while β_1 refers to the standardized CogAT scores. X_i represents demographic characteristics of the students including gender, race, FRL, G/T, and ELL status. μ_{tg} represents grade level and cohort level fixed effects.

To address our second question, we employed Linear Probability Models (LPM) to examine the likelihood of students to score above the 90^{th} and 95^{th} percentile of the CogAT or ACT Aspire tests. LPM is a suitable method to use in this context because it involves linear regression models that are applied to binary outcomes (Chatla & Shmueli, 2016). We assumed that students who scored in the top 5% of state standardized tests were high achievers and could be considered academically talented (e.g., Lakin & Wai, 2020; Wai et al., 2018). Additionally, there are practices to consider scores above the 90^{th} percentile as students who are advanced learners. Therefore, we standardized ACT Aspire and CogAT scores and created a local percentile. Based on this percentile, we created two new pools of students who scored above the 90^{th} and 95^{th} percentile of the CogAT and ACT Aspire tests. We then explored which test's top scoring—above the 90^{th} and 95^{th} percentile of $CogAT_i$ or $ACTAspire_i$— was predictive of greater inclusiveness and diversity. We utilized the following simplified LPM model:

$$Pr(Y_{it}) = \wedge\ (\beta_0 + X_{it} + \varepsilon_{it})$$

The outcome variable Y_{it} is a binary variable that takes the value 1 if student i in year t scored above the 90^{th} or 95^{th} percentile of the CogAT test or 0 if otherwise. Likewise, to examine inclusiveness of the achievement test, variable Y_{it} takes the value 1 if student i in year t scored at or above the 95^{th} percentile of the ACT Aspire test or 0 if otherwise. The variable X_{it} is a matrix that represents students' demographic characteristics, such as their ELL and FRL status, gender, or race/ethnicity.

Findings

We explored the descriptives of both the CogAT and ACT Aspire tests (see Table 2). Across the cohorts, we did not find any noticeable differences in test scores. However, for CogAT Universal Scale Scores (USS scores), we noticed score differences based on demographic differences. For instance, ELL students were the lowest scorers on the CogAT test; their average score was around 12 points lower than the average score. White and Asian students were the highest scorers on the CogAT test. This score difference on the CogAT was consistent across both cohorts in the sample we studied in Arkansas. Whereas, for ACT Aspire test scores, we did not find any noticeable differences (differences of mean scores of more than 2-3 score points) across demographics and cohorts.

Additionally, using locally created percentiles, we identified the top 5% and top 10% scorers on both tests. Table 3 illustrates the distribution of these high-achieving students across different demographic categories for the two cohorts. Overall, a higher percentage of students were in the top 5% and 10% of CogAT scores compared to ACT Aspire scores. In Cohort 1, a larger proportion of male students were represented in the top 5% of both tests, while a higher percentage of female students were in the top 10% of CogAT in Cohort 2. G/T students comprised a substantial proportion of the highest achievers on both assessments across both cohorts, whereas ELL students were consistently the most underrepresented subgroup. White students were the most represented group in the top percentages for both tests and cohorts. The data also showed a higher prevalence of FRL students in the top groups for CogAT compared to the ACT Aspire.

Table 2.
Summary of Test Scores

Variable	Cohort 1						Cohort 2					
	CogAT (USS score)		ACT Aspire (Math Scale Score)		ACT Aspire (ELA Score)		CogAT (USS score)		ACT Aspire (Math Scale Score)		ACT Aspire (ELA Score)	
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*
Overall	197.98	18.49	414.14	4.18	418.66	5.45	197.16	19.39	416.84	5.61	419.21	5.66
Gender												
Female	197.72	17.15	414.08	3.98	419.35	5.37	196.92	18.05	416.96	5.45	420.13	5.39
Male	198.24	19.71	414.2	4.38	417.98	5.44	197.24	19.49	416.21	5.47	419.19	5.54
Educational Characteristics												
ELL	185.87	13.13	411.44	3.74	415.29	4.19	187.02	16.45	412.93	3.84	416.37	4.48
FRL	194.24	16.61	413.25	4.07	417.51	5.06	193.43	18.27	415.21	4.9	417.94	5.45
G/T	223.19	17.44	419.1	2.62	425.45	3.8	213.35	21.21	423.13	4.64	426.16	3.41
Race and Ethnicity												
Asian	202.76	19.51	414.9	4.19	419.7	5.35	203.23	23.87	418.72	6.27	420.16	6.03
Black	191.27	18.08	412.89	4.72	417.71	6.04	195.31	18.47	416.21	5	416.49	5.65
Hispanic	192.96	16.07	413.37	3.98	417.7	4.98	191.37	16.83	414.94	4.59	418.4	5.02
White	202.54	18.96	415.01	4.08	419.65	5.59	200.99	19.77	418.21	5.77	419.7	5.86
Other race	193.94	18.11	412.69	4.3	417.09	5.22	193.29	17.99	413.74	4.39	418.89	5.09

Table 3.
Top 5% and Top 10% of the ACT Aspire and CogAT

Variable	Cohort 1				Cohort 2			
	CogAT		ACT Aspire		CogAT		ACT Aspire	
	Top 5% (%)	Top 10% (%)	Top 5% (%)	Top 10% (%)	Top 5% (%)	Top 10% (%)	Top 5% (%)	Top 10% (%)
Overall	195	448	206	447	387	593	232	451
Gender								
Female	40.00	38.62	48.54	53.02	43.67	45.53	49.14	49.45
Male	60.00	61.38	51.46	46.98	56.33	54.47	50.86	50.55
Educational Characteristics								
ELL	1.54	3.13	1.47	2.7	10.59	9.27	0	0
FRL	29.32	32.6	22.76	29.9	46.25	42.16	28.45	30.82
G/T	76.41	63.84	76.47	67.64	29.72	34.23	78.02	62.75
Race and Ethnicity								
Asian	2.56	2.68	3.43	2.7	4.13	3.71	3.02	3.1
Black	2.05	1.34	2.45	2.02	4.39	3.88	0.86	1.11
Hispanic	20.51	25.00	20.1	24.94	20.93	18.89	17.24	17.29
White	67.69	65.18	70.1	64.04	65.63	69.14	76.72	76.27
Other race	7.18	5.80	3.92	6.29	4.91	4.38	2.16	2.22

Note. Percentages (%) represent the proportion of the total number of students in the "Overall" row who fall into each category.

RQ 1. Association Between CogAT and ACT Aspire Scores

We found a significant positive correlation between ACT Aspire and CogAT scores (Table 4). The overall correlation between ACT Aspire and CogAT scores was $r = .59$. The correlation between these two tests was $r = .72$ for Cohort 1 and $r = .46$ for Cohort 2. We examined the correlation of CogAT scores with ACT Aspire math and ELA scores and found a similar kind of inconsistent association across cohorts. We noticed that, overall, math scale scores and ELA scores had a correlation of $r = .56$ with CogAT scores. For Cohort 1, ACT Aspire math and ELA scores had a correlation of $r = .71$ and $r = .65$, respectively. For Cohort 2, the correlation of CogAT with ACT Aspire math and ELA scores reduced to $r = .46$ and $r = .45$, respectively. Additionally, we tested the correlations between ACT math and ACT ELA scores. The overall correlation was $r = .73$, and it remains consistent across two cohorts.

Regarding subgroups, Table 4 demonstrate a consistent trend of correlation across all the metrics for FRL, ELL, White and non-White students. For example, correlations for FRL students were consistently lower compared to the overall population, while ELL students exhibited the lowest correlations across all correlation categories. We found all the correlations statistically significant.

Table 4.
Correlation Coefficients Between the ACT Aspire and CogAT

Test	Overall	Cohort 1	Cohort 2	FRL	ELL	White	Non-White
ACT Aspire vs CogAT	.59***	.72***	.46***	.52***	.44***	.56***	.57***
ACT Math vs CogAT	.56***	.71***	.46***	.52***	.45***	.55***	.58***
ACT ELA vs CogAT	.56***	.65***	.45***	.49***	.37***	.54***	.53***
ACT Math vs ACT ELA	.73***	.77***	.69***	.71***	.60***	.73***	.71***

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Additionally, we conducted OLS regression analyses, both with and without control variables. The OLS regressions indicated similar associations between these two tests (Table 5). The results of initial OLS regressions without controls revealed a strong association between the two tests. Specifically, a one standard deviation (SD) increase in CogAT scores was associated with an increase of 0.59 SD in ACT Aspire scores ($p < .001$).

However, the association reduced once we included some control variables in the regression. We controlled for demographic and academic variables, i.e., gender, race, FRL, ELL and G/T status. The purpose of including these control variables was to isolate the direct relationship between CogAT and ACT Aspire scores. The inclusion of these controls resulted in a moderated association between the two test scores. We found that a one SD increase in CogAT scores was associated with an increase of 0.39 SD in ACT Aspire scores ($p < .001$), holding all else equal. We observed that factors like students' FRL, ELL and G/T status, and gender explained significant variations in the relationship between these two tests ($p < .001$), suggesting that the scores on achievement tests and ability tests may vary depending on students' demographic and socio-economic status.

Next, we extended our analysis to examine the association between these two tests across different cohorts, revealing notable variations as we saw in the correlation results (Table 4). For Cohort 1, we noticed that a one SD increase in CogAT scores was associated with an increase of 0.57 SD in ACT Aspire scores ($p < .001$), holding all else equal. In contrast, Cohort 2 exhibited a weaker association, with a one SD increase in CogAT associated with an increase of 0.27 SD in ACT Aspire scores ($p < .001$), holding all else equal. These cohort-specific findings underscore the potential variability in the relationship between cognitive ability and academic achievement measures across different students groups, at least within the samples in our study.

Table 5.
Results of linear regressions for associations between the ACT Aspire and CogAT tests

			ACT Aspire		ACT Aspire (Cohort 1)		ACT Aspire (Cohort 2)	
	β	***SE***	***β***	***SE***	***β***	***SE***	***β***	***SE***
CogAT	.59***	0.01	.39***	0.01	.57***	0.01	.27***	0.01
Demographics								
Female			.07***	0.02	.15***	0.02	.02	0.02
Asian			.12*	0.05	.16*	0.08	.04	0.10
Black			-.08	0.05	.08	0.07	-.16*	0.07
Hispanic			.15***	0.02	.32***	0.03	-.01	0.03
Other race			.03	0.03	.01	0.04	.01	0.04
Programmatic Status								
FRL			-.26***	0.02	-.26***	0.03	-.24***	0.03
ELL			-.45***	0.02	-.40***	0.04	-.48***	0.03
G/T			.78***	0.03	.42***	0.04	.96***	0.03
Constant	.04	0.01	.09***	0.02	.02	0.02	.13***	0.02
N	10,338		8,828		3,767		5,061	
R^2	0.35		0.43		0.55		0.39	

Note. Standard errors are robust. Comparison group for the races (Asian, Black, Hispanic and other races) were White.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

RQ2. Predictors of Greater Diversity and Inclusiveness

Next, we examined aspects of diversity and inclusiveness under both the ACT Aspire and CogAT tests. Using a conservative approach to identification, we tested students' likelihood of scoring above the 90th and 95th percentile of the ACT Aspire and CogAT tests based on their FRL and ELL status, race and gender. For race, we used 'White' as the reference category.

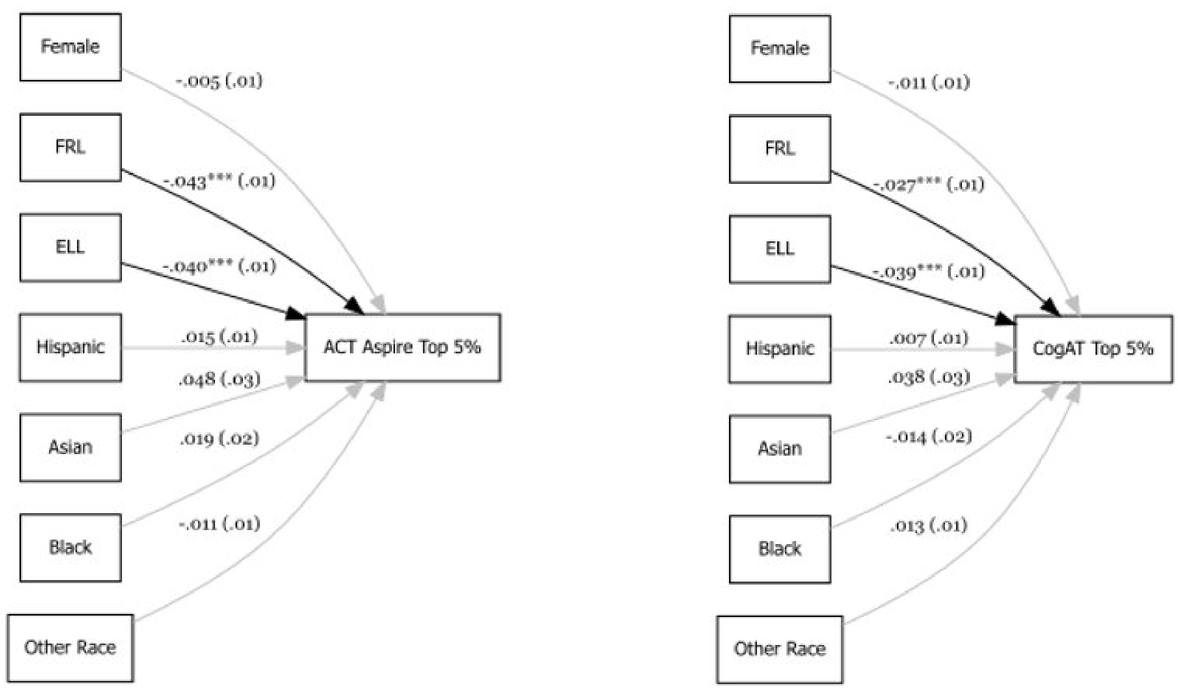

Figure 1. Likelihood of Cohort 1 subgroups to score in the top 5% on the CogAT and ACT Aspire tests.

Note. Robust standard errors are in parentheses. Coefficients of Black, Hispanic, Asian and other races are drawn compared to White as the base or reference category. Bold paths indicate a significant association, while gray paths indicate a path that was not significantly different between the groups.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

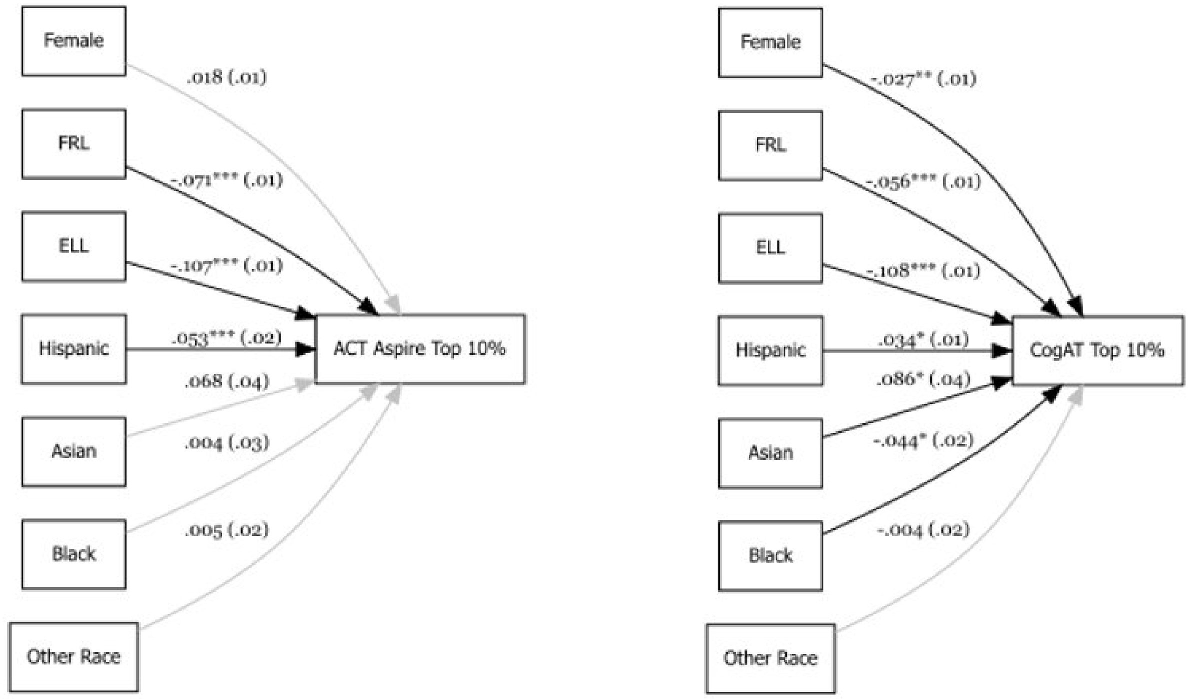

Figure 2. Likelihood of Cohort 1 subgroups to score in the top 10% on the CogAT and ACT Aspire tests.

Note. Robust standard errors are in parentheses. Coefficients of Black, Hispanic, Asian and other races are drawn compared to White as base or reference category. Bold paths indicate a significant association, while gray paths indicate a path that was not significantly different between the groups.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

We first considered the students who scored above the 95th percentile on ACT Aspire and CogAT scores. Our analysis revealed intriguing patterns across cohorts. In Cohort 1, we observed no statistically significant differences based on gender and race/ethnicity for either test (Figure 1). However, FRL and ELL students demonstrated a lower likelihood of achieving top 5% scores compared to their non-FRL and non-ELL counterparts. For example, our LPM results indicated that, all else equal, FRL students compared to non-FRL students, in Cohort 1, were four percentage points (pp) and three pp less likely to score above the 95th percentile of the ACT Aspire and CogAT, respectively ($p < .001$). Similarly, ELL students, compared to non-ELL students, in Cohort 1, were four pp less likely to achieve this benchmark on both tests ($p < .001$), holding all else equal.

Cohort 2 presented a more nuanced picture. For ACT Aspire top 5%, in addition to the disparities observed for FRL and ELL students, Figure 3 demonstrated that Black students were four pp less likely than White students to score above the 95th percentile ($p < .001$). The CogAT top 5% results for this cohort revealed further disparities, with female students, Hispanic students and students of other races showing a lower likelihood to score in the top 5% of the CogAT (see Figure 3) compared to their respective counterparts. The significance level for these results ranged from 95% to 99%.

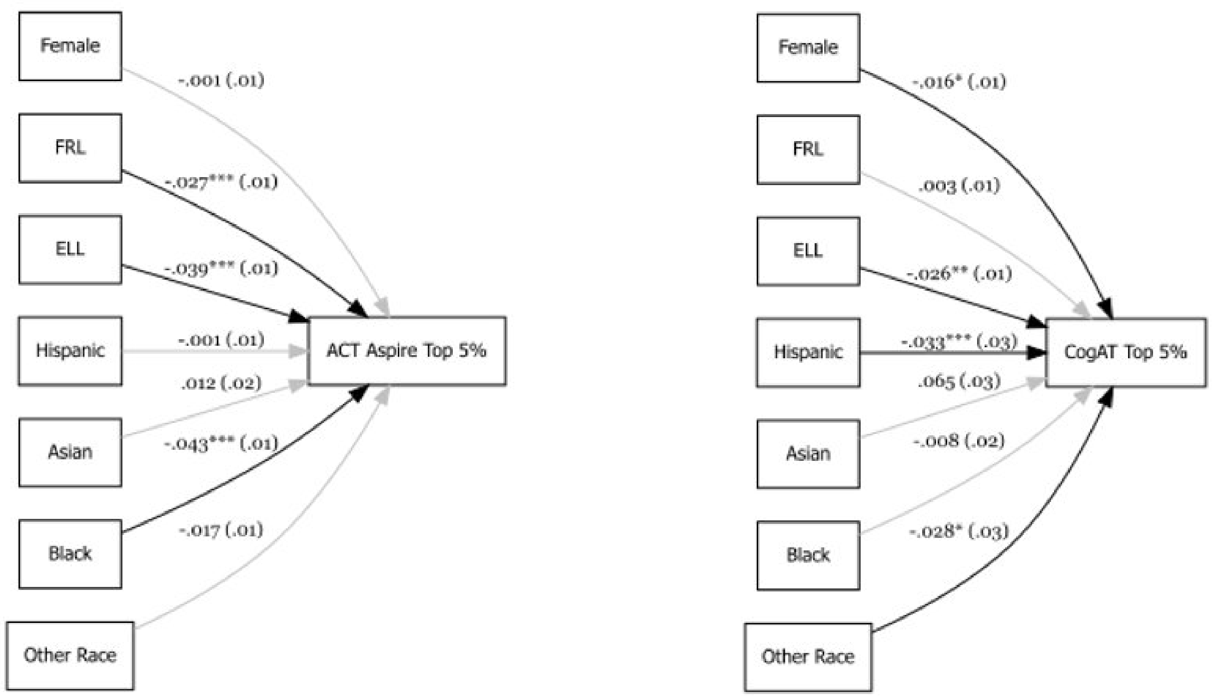

Figure 3. Likelihood of Cohort 2 subgroups to score in the top 5% on the CogAT and ACT Aspire tests.

Note. Robust standard errors are in parentheses. Coefficients of Black, Hispanic, Asian and other races are drawn compared to White as base or reference category. Bold paths indicate a significant association, while gray paths indicate a path that was not significantly different between the groups.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

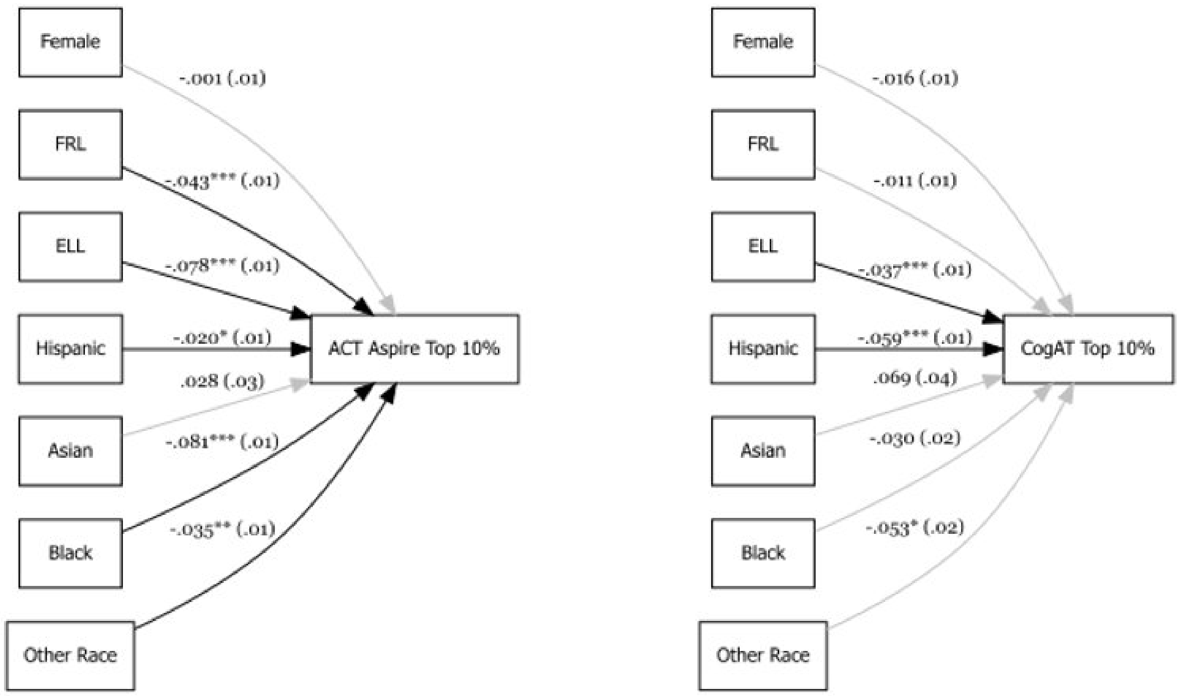

Figure 4. Likelihood of Cohort 2 subgroups to score in the top 10% on the CogAT and ACT Aspire tests.

Note. Robust standard errors are in parentheses. Coefficients of Black, Hispanic, Asian and other races are drawn compared to White as base or reference category. Bold paths indicate a significant association, while gray paths indicate a path that was not significantly different between the groups.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Next, considering the students who scored above the 90th percentile on the ACT Aspire and CogAT tests, we uncovered distinct patterns between the two tests. In Cohort 1, the ACT Aspire showed statistically significant differences only for FRL, ELL and Hispanic students. FRL and ELL students were less likely to score above the 10th percentile on the ACT Aspire relative to non-FRL and non-ELL peers ($p < .001$), while Hispanic students showed a higher likelihood of doing so relative to White students ($p < .001$).

The CogAT results for Cohort 1's top 10%, on the other hand, revealed more widespread disparities. All the subgroups, except for students from "other races" showed statistically significant group differences. Female, FRL, ELL and Black students appeared to be less likely to score above the 90th percentile on the CogAT test, compared to their respective counterparts. Interestingly, Hispanic students were more likely to score above the 90th percentile compared to White students, mirroring the ACT Aspire results.

Cohort 2 presented a noticeably different landscape for top 10% scorers. For the ACT Aspire, all subgroups except Asian and female students showed statistically significant differences, with FRL, ELL, Hispanic, Black, and other race students less likely to score above the 90th percentile on the ACT Aspire test. Contrastingly, CogAT results for this cohort showed a statistically significant difference only for ELL and Hispanic students.

Discussion and Conclusion

This study examined the extent to which CogAT and ACT Aspire scores were associated and the nature of diversity and inclusiveness that both tests offer in the two samples of roughly 5,000 each in Arkansas that were examined. Two cohorts of students drawn from the years 2018 through 2022 were included in the sample. Although this study does not show any causal claims, it offers useful information and insights to stakeholders such as state policymakers, education researchers, G/T counselors, G/T teachers and school leaders, especially those in Arkansas.

Use of the ACT Aspire and CogAT

The findings of this study demonstrate that there is a considerable association between the ACT Aspire and CogAT tests, which confirms the long history of all cognitive and achievement tests showing substantial positive correlations (Wai et al., 2018). This result confirms findings from Naglieri et al. (2003), Peng and Kievit (2020) and Zins and Barnett (1983) that achievement and reasoning tests tend to be highly correlated. Our finding about the significant correlation between these two tests suggests that maybe it is reasonable to use the CogAT or ACT Aspire tests interchangeably as an objective indicator to identify G/T learners.

However, there are two notable concerns raised from the findings that pose alternative interpretations and related policy. First, our results demonstrated a noticeable variation of correlations across cohorts; Cohort 1 portrayed a correlation of $r = .71$ while Cohort 2's correlation was $r = .46$—a difference of $r = .25$. We used Fisher's test and found this difference in correlations between the two cohorts was statistically significant. Similar differences arose in the OLS regressions as well. For our second research question, we found significant differences between cohorts in predicting diversity among the students scoring above the 95th and 90th percentile.

This discrepancy between these two cohorts warrants further investigation, particularly for Cohort 2. However, Table 4 highlights a consistent pattern of correlation between ACT Aspire Math and ACT Aspire ELA scores across cohorts (Overall $r = .73$, Cohort 1 $r = .77$, Cohort 2 $r = 0.69$). This consistency aligns with the findings of Wai et al. (2009) as well as others who study math and verbal achievement/ability tests, who reported a widely replicated math-verbal correlation of approximately $r = .76$ found in population-representative samples, such as Project Talent. This alignment suggests that the observed results in our Cohort 2 sample are not entirely anomalous, though the lowered correlation between math and verbal test scores suggest that these results may be atypical and Cohort 1 findings may be more robust and aligned with prior research.

These variations and inconsistencies across cohorts, however, do suggest more caution to the idea of the interchangeable use of achievement tests and ability tests. Though this may also be a reflection of the unique variation in our samples studied, rather than generalizable findings for policy.

Second, the OLS regression results (see Table 5) showed that the coefficients of association between the two tests reduced significantly when we added control variables. The coefficient was .59 without any control variable but reduced to .39 once control variables were added. However, these variations due to control variables were predictable, as the correlation coefficients for subgroups (see Table 4) demonstrated variability. This finding is also consistent with Wai et al. (2009) who observed stable associations of correlations between low SES and math, verbal, or spatial scores in population-level samples. These variations of association between the two tests due to control variables suggest that the scores on achievement test and ability tests and their potential correlation may vary depending on students' demographic and socio-economic status. Therefore, G/T coordinators and teachers may need to be more careful in their use of achievement test scores and ability test scores. But again, using what is available in a universal screening context may be better than not using a universally applied objective indicator at all. Thus, a counterargument would be that using one of these tests in practice as a first universal screener (typically the achievement test) is a reasonable step, but that also including another ability test or additional tests would be even better.

Diversity and Inclusiveness

The LPM results (see Figures 1-4) underscore the complex interplay of demographic factors in standardized test performance in our samples. The likelihood of high achievement on the ACT Aspire and CogAT tests varies across cohorts and tests for students from different socio-economic and racial/ethnic backgrounds. This consistent variability between cohorts and tests suggests that the relationships between student characteristics and test performance is influenced by factors beyond individual student attributes. These results call for a nuanced approach to G/T program identification, one that considers multiple measures and accounts for the diverse backgrounds and experiences of students.

The recommendation of considering scores from multiple tests as an objective indicator, which is universally used, reiterates the findings of Ozen et al. (2024), and is ideal in an optimal situation, but is not the typical situation in practice. Moreover, we already know the benefits of universal screening: many findings in the past in the tests/selection literature and the more recent findings of Card and Giuliano (2016) reveal that testing all students leads to a significant increase in the representation of low-income and historically underserved students in gifted programs. Also, universal screening can help address, at least in part, the systemic failure to recognize the potential of financially disadvantaged students, and can also enhance diversity in gifted programs (Callahan et al., 2013; Wai & Lakin, 2020).

Additionally, the variations in the pathways to the top 5% and 10% for different subgroups of students across the tests and cohorts suggest that opting for only the students above the 95th percentile on the ability test (CogAT) would exclude a considerable number of students with the highest scores on their achievement test (ACT Aspire). Whereas, using both tests in considering students for G/T programs would offer more high-achieving students a unique opportunity to receive the benefits of the G/T services. This example aligns with what Lohman (2005) has long discussed, so in a way it is not new. Lohman (2005), using samples from Naglieri and Ronning (2000), showed that selecting the top 5% on the ability test in the G/T program would identify only 31% of the students in the top 5% of the math achievement test, excluding 69% of the students with the best mathematics achievement, suggesting the importance of using more than one test score in G/T identification, in addition to using multiple factors in the identification process to help identify more diverse talent.

Limitations and Future Research

According to Tran et al. (2022), as much as 30% of the students in top 5% on both third-grade literacy and math were not identified as gifted in the Arkansas sample they studied. The Civil Rights (2018) report and findings of Gentry et al. (2019) also reiterated these findings regarding the underrepresentation of students from disadvantaged communities. We believe that the findings of this study: use of multiple test scores—achievement and ability—as a universal screener to widen the net of G/T identification can be useful potential tools to address the existing disparity in the G/T identification process. However, the unique demographic nature of the school districts studied may attenuate the external validity of this study (our study is specific to our samples and our locale, and is not population representative). This study also used samples that were assessed before and after the pandemic, thus there may have been learning loss related attenuation of typical correlations between measures due to this important historical event. Also, since we did not use school or district-level modeling, perhaps parts of our suggestions about the local norms are not robustly supported by our analysis. Therefore, this study may continue with richer data, more robust models and newer exam formats to find out justify the findings of this study in order to ensure universal and local norm-based screening for G/T identification process. When connected with other studies, perhaps a larger more consistent and replicated set of findings for policy may emerge.

Additionally, we do not know what the exact objective and subjective measures are that G/T coordinators use when they consider G/T placement and how these practices vary across schools and school districts, not only within the state across districts, but more broadly in all kinds of G/T identification procedures. For example, in the state of Arkansas, one of the objective indicators used in identification must be a creativity measure, and we were not able to assess that in our study. Future studies that explore the unique perceptions and practices of the practitioners in identifying G/T learners in our local context (or other local contexts) could provide valuable insights for the policymakers and researchers.

Our data has additional limitations, as it does not include information on students' scores for individual CogAT test batteries. Access to this detailed data would enable a more precise comparison between corresponding components of the CogAT and ACT Aspire tests. Future research investigating the associations between similar test components could significantly enhance the literature in this area. Finally, we reiterate that our findings are based on relatively smaller samples of data in Arkansas, and may not necessarily generalize to other states or contexts regarding policy decisions, such as at a different point in time.

References

Acar, S., Sen, S., & Cayirdag, N. (2016). Consistency of the performance and nonperformance methods in gifted identification: A multilevel meta-analytic review. *Gifted Child Quarterly*, *60*(2), 81-101. https://doi.org/10.1177/0016986216634438

ACT Aspire. (n.d.). ACT. Retrieved December 8, 2024, from https://global.act.org/content/global/en/products-and-services/act-aspire.html

Assouline, S. G., Colangelo, N., VanTassel-Baska, J., & Lupkowski-Shoplik, A. E. (Eds.). (2015). *A nation empowered: Evidence trumps the excuses that hold back America's brightest students*. Iowa City, IA: The Belin-Blank Center for Gifted and Talented Education.

Callahan, C. M. (2005). Identifying gifted students from underrepresented populations. *Theory Into Practice, 44*, 98-104. 10.1207/s15430421tip4402_4

Callahan C. M., Moon, T. R., & Oh, S. (2013). *Status of elementary gifted programs*. Charlottesville: University of Virginia, The National Research Center on the Gifted and Talented. http://www.nagc.org/resources-publications/resources-university-professionals

Card, D., & Giuliano, L. (2016). Universal screening increases the representation of low-income and minority students in gifted education. *Proceedings of the National Academy of Sciences*, *113*(48), 13678-13683. https://doi.org/10.1073/pnas.1605043113

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*(2), 193-209. https://doi.org/10.1177/0016986217752097

Chatla, S., & Shmueli, G. (2016). *Linear Probability Models (LPM) and big data: The good, the bad, and the ugly* (SSRN Scholarly Paper 2353841). https://doi.org/10.2139/ssrn.2353841

Dynarski, S. M. (2018). *ACT/SAT for all: A cheap, effective way to narrow income gaps in college.* Washington, DC: Brookings Institution. https://www.brookings.edu/articles/act-sat-for-all-a-cheap-effective-way-to-narrow-income-gaps-in-college/

Gentry, M. (2009). Myth 11: A comprehensive continuum of gifted education and talent development services: Discovering, developing, and enhancing young people's gifts and talents. *Gifted Child Quarterly*, *53*(4), 262-265. https://doi.org/10.1177/0016986209346937

Giessman, J. A., Gambrell, J. L., & Stebbins, M. S. (2013). Minority performance on the Naglieri Nonverbal Ability Test, versus the Cognitive Abilities Test, Form 6: One gifted program's experience. *Gifted Child Quarterly, 57*, 101-109. 10.1177/0016986213477190

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

Gómez-Veiga, I., Vila Chaves, J. O., Duque, G., & García Madruga, J. A. (2018). A new look to a classic issue: Reasoning and academic achievement at secondary school. *Frontiers in Psychology*, *9*, 400. https://doi.org/10.3389/fpsyg.2018.00400

Harradine, C. C., Coleman, M. R. B., & Winn, D. C. (2014). Recognizing academic potential in students of color: Findings of U-STARS~PLUS. *Gifted Child Quarterly, 58*, 24-34. 10.1177/0016986213506040

Hodges, J., Tay, J., Maeda, Y., & Gentry, M. (2018). A meta-analysis of gifted and talented identification practices. *Gifted Child Quarterly*, *62*(2), 147-174. https://doi.org/10.1177/0016986217752107

Hoxby, C. M., & Avery, C. (2013, Spring). The missing "one-offs": The hidden supply of high-achieving, low-income students. *Brookings Papers on Economic Activity*, *2013(1)*, 1-65. 10.3386/w18586

Lakin, J. M. (2019). *Universal screening in gifted and talented identification: Implementation and overcoming challenges*. Retrieved September 4, 2024, from https://blog.riversideinsights.com/universal-screening-in-gifted-and-talented-identification

Lakin, J. M. (2016). Universal screening and the representation of historically underrepresented minority students in gifted education: Minding the gaps in Card and Giuliano's research. *Journal of Advanced Academics, 27*(2), 139-149. https://doi.org/10.1177/1932202X16630348

Lakin, J. M., & Lohman, D. F. (2011). The predictive accuracy of verbal, quantitative, and nonverbal reasoning tests: Consequences for talent identification and program diversity. *Journal for the Education of the Gifted, 34*(4), 595-623. https://doi.org/10.1177/016235321103400404

Lakin, J. M., & Wai, J. (2020). Spatially gifted, academically inconvenienced: Spatially talented students experience less academic engagement and more behavioural issues than other talented students. *British Journal of Educational Psychology, 90*(4), 1015-1038. https://doi.org/10.1111/bjep.12343

Lohman, D. F. (2005). The role of nonverbal ability tests in identifying academically gifted students: an aptitude perspective. *Gifted Child Quarterly*, *49*(2), 111-138. https://doi.org/10.1177/001698620504900203

Lohman, D. F., & Lakin, J. M. (2010). Consistencies in sex differences on the Cognitive Abilities Test across countries, grades, test forms, and cohorts. *British Journal of Educational Psychology*, *79*(2), 389–407. https://doi.org/10.1348/000709908X354609

Lubinski, D., & Benbow, C. P. (2006). Study of Mathematically Precocious Youth after 35 years: Uncovering antecedents for the development of math-science expertise. *Perspectives on Psychological Science, 1*(4), 316-345. https://doi.org/10.1111/j.1745-6916.2006.00019.x

Lubinski, D. & Benbow, C. P. (2000). States of excellence. *American Psychologist, 55*(1), 137-150. 10.1037//0003-066x.55.1.137

May, H., Johnson, I. P., Haimson, J., Sattar, S., & Gleason, P. (2009). Using state tests in education experiments: A discussion of the issues (NCEE 2009–013). Washington, DC: National Center for Education Evaluation and Regional Assistance.

McBee, M. T. (2010). Examining the probability of identification for gifted programs for students in Georgia elementary schools: A multilevel path analysis study. *Gifted Child Quarterly*, *54*, 283-297. 10.1177/0016986210377927

McBee, M. T. (2006). A descriptive analysis of referral sources for gifted identification screening by race and socioeconomic status. *Journal of Secondary Gifted Education*, *17*(2), 103-111. https://doi.org/10.4219/jsge-2006-686

McBee, M. T., & Makel, M. C. (2019). The quantitative implications of definitions of giftedness. *AERA Open*, *5*(1). https://doi.org/10.1177/2332858419831007

Naglieri, J. A., & Ford, D. Y. (2003). Addressing underrepresentation of gifted minority children using the Naglieri Nonverbal Ability Test (NNAT). *Gifted Child Quarterly*, *47*(2), 155-160. https://doi.org/10.1177/001698620304700206

Naglieri, J. A., & Ronning, M. E. (2000). The relationship between general ability using the Naglieri Nonverbal Ability Test (NNAT) and Stanford Achievement Test (SAT) reading achievement. *Journal of Psychoeducational Assessment, 18,* 230-239. https://doi.org/10.1177/073428290001800303

Olszewski-Kubilius, P., & Clarenbach, J. (2012). *Unlocking emergent talent: Supporting high achievement of low-income, high ability students*. Washington, DC: National Association for Gifted Children. https://eric.ed.gov/?id=ED537321

Ozen, Z., Pereira, N., Karatas, T., Castillo-Hermosilla, H., & Maeda, Y. (2025). A meta-analytic evaluation: Investigating evidence for the validity of the Cognitive Abilities Test. *Gifted Child Quarterly*, *69*(1), 3-15. https://doi.org/10.1177/00169862241285593

Peng, P., & Kievit, R. A. (2020). The development of academic achievement and cognitive abilities: A bidirectional perspective. *Child Development Perspectives, 14*(1), 15-20. https://doi.org/10.1111/cdep.12352

Peters, S. J., & Carter, J. S. (2022). Predictors of access to gifted education: What makes for a successful school? *Exceptional Children*, *88*(4), 341-358. https://doi.org/10.1177/00144029221081092

Peters, S. J., Makel, M. C., & Hernandez, K. R. (2021). Local norms for gifted and talented student identification: Everything you need to know. *Gifted Child Today*, *44*(2), 93-104. 10.1177/1076217520985181

Peters, S. J., Gentry, M., Whiting, G. W., & McBee, M. T. (2019). Who gets served in gifted education? Demographic representation and a call for action. *Gifted Child Quarterly*, *63*(4), 273-287. https://doi.org/10.1177/0016986219833738

Peters, S. J., Rambo-Hernandez, K. R., Makel, M. C., Matthews, M. S., & Plucker, J. A. (2019). Effects of local norms on racial and ethnic representation in gifted education. *AERA Open, 5*(2), 1-18. https://doi.org/10.1177/2332858419848446

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

Rasheed, M. (2020). Context and content in rural gifted education: A literature review. *Journal of Advanced Academics*, *31*(1), 61-84. https://doi.org/10.1177/1932202X19879174

Reis, S. M., & Renzulli, J. S. (2010). Is there still a need for gifted education? An examination of current research. *Learning and Individual Differences*, *20*(4), 308-317. https://doi.org/10.1016/j.lindif.2009.10.012

Reis, S. M., & Renzulli, J. S. (1991). The assessment of creative products in programs for gifted and talented students. *Gifted Child Quarterly*, *35*(3), 128-134. https://doi.org/10.1177/001698629103500304

Renzulli, J. S. (2005). The Three-Ring conception of giftedness: A developmental model for promoting creative productivity. In R. J. Sternberg & J. E. Davidson (Eds.), *Conceptions of giftedness* (pp. 246-279). Cambridge University Press.

Subotnik, R. F., Olszewski-Kubilius, P., & Worrell, F. C. (2011). Rethinking giftedness and gifted education: A proposed direction forward based on psychological science. *Psychological Science in the Public Interest*, *12*(1), 3-54. https://doi.org/10.1177/1529100611418056

Sussman, J., & Wilson, M. R. (2019). The use and validity of standardized achievement tests for evaluating new curricular interventions in mathematics and science. *American Journal of Evaluation*, *40*(2), 190-213. https://doi.org/10.1177/1098214018767313

Tran, B., McKenzie, S. C., & Wai, J. (2022). Local norms and gifted and talented identification in Arkansas: Can it help improve student diversity? *Office for Education Policy Arkansas Education Report, 19*(7), 1-23. https://scholarworks.uark.edu/oepreport/90

Tran, B., Wai, J., McKenzie, S. C., Mills, J. & Seaton, D. (2021). Expanding gifted identification to capture academically advanced, low-income, or other disadvantaged students: The case of Arkansas. *Journal for the Education of the Gifted*, *45*(1), 64-83. 10.1177/01623532211063936

U.S. Department of Education, Office for Civil Rights. (2018). Civil Rights Data Collection (CRDC).

Wai, J., Brown, M. I., & Chabris, C. F. (2018). Using standardized test scores to include general cognitive ability in education research and policy. *Journal of Intelligence, 6*(3), 37. https://doi.org/10.3390/jintelligence6030037

Wai, J., & Lakin, J. M. (2020). Finding the missing Einsteins: Expanding the breadth of cognitive and noncognitive measures used in academic services. *Contemporary Educational Psychology, 63,* 101920. https://doi.org/10.1016/j.cedpsych.2020.101920

Wai, J., Lubinski, D., Benbow, C. P., & Steiger, J. H. (2010). Accomplishment in science, technology, engineering, and mathematics (STEM) and its relation to STEM educational dose: A 25-year longitudinal study. *Journal of Educational Psychology, 102*(4), 860-871. https://doi.org/10.1037/a0019454

Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology, 101*(4), 817–835. https://doi.org/10.1037/a0016127

Wai, J., & Worrell, F. C. (2020). How talented low-income kids are left behind. *Phi Delta Kappan, 102*(4), 26-29. https://doi.org/10.1177/0031721720978058

Williamson, J. (2019). Predictive ability of ACT Aspire for ACT at Star City high school. *Theses and Dissertations from 2019*. 11. https://orc.library.atu.edu/etds_2019/11

Wyner, J. S., Bridgeland, J. M., & DiIulio, J. J. (2007). *Achievement trap: How America is failing millions of high-achieving students from lower-income families*. Jack Kent Cooke Foundation.

Warne, R. T., & Larsen, R. A. A. (2022). School integration limits the ability of local norms to diversify gifted Programs: A mathematical analysis with implications related to the achievement gap. *Journal of Advanced Academics*, *33*(2), 275-309. https://doi.org/10.1177/1932202X211069078