

Poor Results for



High Achievers

For nearly a decade, the No Child Left Behind Act (NCLB) has focused the attention of policymakers and researchers squarely on the achievement of low-performing students, with some apparent success. The math and reading scores on the National Assessment of Educational Progress of the nation's lowest-achieving 10 percent of 4th and 8th graders have risen sharply since 2000, continuing a trend that began in the 1990s. Yet some may wonder about the potential cost of this focus on higher-achieving students, for whom improvements over the same time period have been modest. Among the questions related to this debate is whether additional programs and resources should be devoted to students on the higher end of the spectrum, those considered gifted.

Three million students in the United States are classified as gifted, yet little is known about the effectiveness of traditional gifted and talented (G&T) programs. In theory, G&T programs might help high-achieving students because they group them with other high achievers and typically offer specially trained teachers and a more advanced curriculum. While previous research indicates that ability grouping is in fact correlated with higher achievement, these findings could be misleading if students placed in high-ability classrooms were likely to be successful for reasons that researchers are unable to measure, such as stronger motivation. To our knowledge, no existing studies

offer convincing evidence on the causal effect of G&T programs on student achievement.

Our research begins to fill this gap with two studies of the G&T programs available to high-achieving middle-school students in a large urban school district in the southwestern United States which, to preserve anonymity we shall refer to as LUSD. Since 2007, all 5th-grade students in LUSD have been evaluated to determine eligibility for gifted and talented programs starting in 6th grade. Those students who are deemed eligible often are grouped in classes with other gifted students. They are also permitted to apply for admission to two middle schools that have oversubscribed magnet G&T programs.

The two studies use different methods to ask distinct but closely related questions. The first exploits the fact that eligibility for G&T programming in LUSD is determined by a well-defined cutoff in students' evaluation scores. By comparing students who score just above the cutoff to those who score just

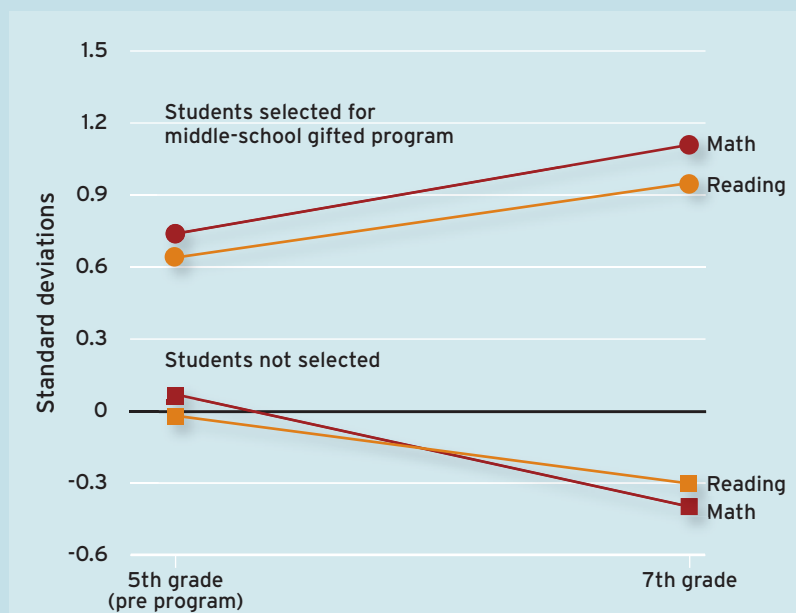
below, the study provides evidence on the effect of enrollment in a G&T program on achievement for those students on the margin of eligibility. The second study takes advantage of the randomized lotteries that determine admission to the district's two premier magnet G&T programs. By comparing students who win the lottery and attend the magnet G&T schools to those who lose the lottery and attend other "neighborhood" programs, the

New evidence on the impact of gifted and talented programs

By SA BUI, STEVEN CRAIG, AND SCOTT IMBERMAN

Chosen Few (Figure 1)

Gifted programs pick good students who are getting better.



Note: Students are classified as gifted/nongifted based on enrollment in 7th grade. Student achievement is standardized separately by grade and subject to have a mean of zero and a standard deviation of 1.

SOURCE: Author's calculations

research provides evidence on whether the magnet G&T programs provide any additional benefits.

The results of both studies will be discouraging for those hopeful that current G&T programs provide a means to accelerate the progress of our most capable students. The first shows that barely eligible students who participated in LUSD's G&T curriculum for all of 6th grade and half of 7th grade exhibit no significant improvement in test scores across a range of subjects, despite their being surrounded by higher-achieving peers and taking more advanced courses. The lottery study corroborates these results, as students admitted to the G&T magnet schools show little improvement in test scores by 7th grade, despite having higher-achieving peers and being taught by more effective teachers. The lone exception is in science, where students admitted to G&T magnet schools performed at substantially higher levels.

It is difficult to know what accounts for these puzzling results. Our best guess, which we discuss in detail below, is that being placed with higher-achieving peers is not all that it is cracked up to be. Students admitted to both types of G&T programs suffer a large drop in their relative rank in terms of grades within their classes, which could have adverse consequences that offset

any benefits of improvements in their educational environment. But we are getting ahead of ourselves. Let's first take a closer look at the programs and the evidence on their effects.

Gifted Students in LUSD

LUSD is a large school district, with more than 200,000 students. The district is heavily minority and very low income; the minority population is more heavily Hispanic than African American. All LUSD students are evaluated for placement in middle-school G&T programs during 5th grade, including those who participated in the district's G&T program in elementary school. In order to be deemed eligible for the middle school G&T program, a student must meet the eligibility criteria set forth in the "gifted and talented identification matrix." The matrix converts scores on standardized tests—the Stanford Achievement Test for English-speaking students and the Aprenda exam for Spanish-speaking students with limited English proficiency—scores on the Naglieri Nonverbal Ability Test (NNAT), average course grades, teacher recommendations, and indicators for socioeconomic status into an overall index score.

While all students who meet these requirements qualify for the G&T program, not all end up being classified as G&T, because parents are allowed to opt out. Some students also enroll in the program initially but later withdraw. Schools in LUSD have a monetary incentive for attracting gifted students, as LUSD provides a funding boost of 12 percent over the average allotment for a regular student.

Gifted students in LUSD are far less likely to be economically disadvantaged and more likely to be white or Asian than other students in the district. They also perform at far higher levels on the Stanford Achievement Tests, which the district administers annually in five subjects: math, reading, language, social science, and science. Their advantage in math and reading test scores in 5th grade is roughly 0.7 of a standard deviation, which amounts to well over two years of academic progress (see Figure 1). By the time the same students have reached 7th grade, these gaps have widened to 1.5 standard deviations in math and 1.25 standard deviations in reading. While this pattern suggests that the students enrolled in the district's G&T programs learn at a faster rate between 5th and 7th grade, it does not necessarily mean that the G&T programs are the cause. It is to that question we now turn.

Effects on Barely Eligible Students

Our first study examines the effects of participation in a G&T program on students who were just barely eligible to participate based on their overall index scores. We focus on students who were evaluated for G&T eligibility as 5th graders in the spring of 2008 for whom we are able to observe outcomes as 7th graders in the 2009–10 school year. Our outcome measures include Stanford Achievement Test scores and attendance rates, both of which are drawn from administrative data provided by the district. After restricting the sample to students near the G&T eligibility cutoff, we are able to examine these outcomes for roughly 2,600 students.

The method used in the study, known as regression discontinuity analysis, takes advantage of the fact that the district uses a strict numerical cutoff in the index score assigned to students as 5th graders in order to determine their eligibility to participate in the G&T program the following year. Because the students are unable to precisely manipulate their index scores, those scoring just below the eligibility cutoff should be very similar to those scoring just above the cutoff. We can therefore attribute any differences in student outcomes on either side of the cutoff to the effect of having being deemed eligible.

As noted above, not all eligible students end up participating in G&T programs due to factors such as a parent’s decision to opt out. Similarly, some students who do not initially qualify later become eligible through an appeals process that allows parents to submit an alternative standardized test score or through additional evaluations conducted in 6th grade. As a result, we use standard statistical techniques to account for the fact that the cutoff our regression discontinuity analysis exploits is “fuzzy” rather than sharp. This allows us to provide evidence on the effects of actual participation in the G&T program, not simply eligibility for it.

Before looking at student outcomes, we first used the same method to confirm that participation in the district’s standard G&T programs led to measurable differences in students’ educational experiences. Clearly, it did. The average achievement of the peers in G&T students’ classrooms were between 0.25 and 0.33 of a standard deviation higher in each core academic subject. Participation in the G&T program also increased the number of advanced courses in which students enrolled in 6th and 7th grade. We found no evidence, however, that the teachers to whom students in the G&T program were assigned were any more effective, as measured by their impact on student test scores.

Did these improvements in peer characteristics and curricular rigor translate into improved outcomes? Our results



indicate that they did not (see Figure 2). Our estimates of the effects of G&T participation for barely eligible students are close to zero in all five subjects and are sufficiently precise to allow us to rule out with 90 percent confidence effects as small as 0.04 standard deviations (sd) in math, 0.07 sd in reading,

Being placed with higher-achieving peers is not all that it is cracked up to be. Students admitted to G&T programs suffer a large drop in their relative rank in terms of grades within their classes.

0.12 sd in language, 0.10 sd in social studies, and 0.19 sd in science. We also looked at the impact of G&T participation for specific student subgroups defined by gender, race/ethnicity, socioeconomic status, and whether the students had been classified as gifted in elementary school. We found little evidence of differential impacts for students in any of these groups.

The Effects of G&T Magnet Programs

Why does the G&T program in LUSD not yield benefits for students on the margin of eligibility? One reason could

be that the qualification boundary is set so low that such students are not able to take advantage of the programs' purported benefits. Our second analysis, which uses experimental research methods to study the effects of enrollment in the district's G&T magnet programs, is intended to shed light on this concern.

LUSD has 41 middle schools, of which 8 have G&T magnet programs, and 2 of these are oversubscribed. As a result, the district uses lotteries to determine which students will be admitted as 6th graders. Our analysis compares the performance of students who win the lottery and attend one of the G&T magnet programs to those who lose the lottery and either attend a neighborhood G&T program in the district, a magnet school based on a different specialty, or a charter school. Because the lottery is random, any differences in outcomes between lottery winners and losers can be attributed to the effect of enrolling in the G&T magnet program rather than one of these alternatives. Moreover, the results of this analysis will apply to the entire population of students who chose to apply.

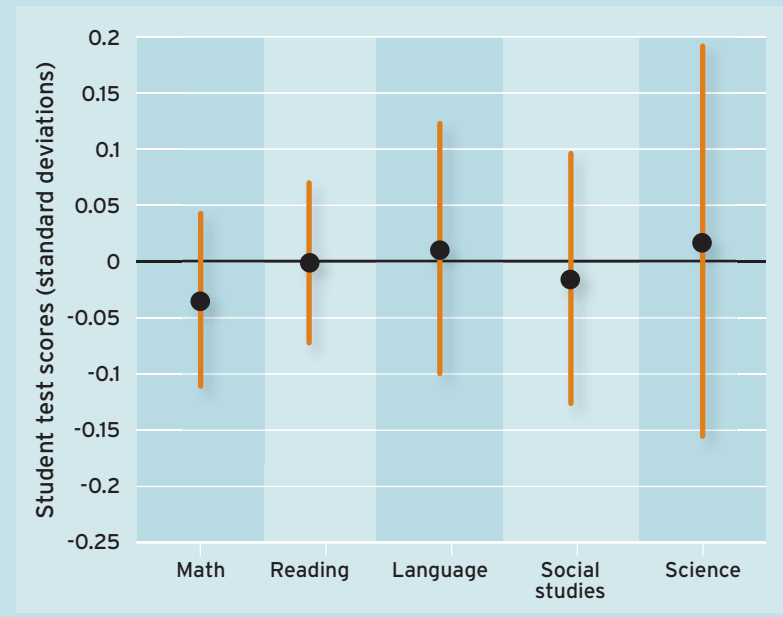
Our lottery analysis is based on the sample of LUSD 5th-grade students determined to be eligible for G&T programs in 2007–08 who applied for admission to one of the two middle schools with an oversubscribed G&T magnet program. This group includes 542 students, 394 of whom were offered admission and 148 of whom were not. We find no statistically significant differences in the observed characteristics of lottery winners and losers, suggesting that the lotteries were in fact conducted in a random way.

Our results provide little evidence that attending a G&T magnet program leads to improvements in student achievement. The one exception is science test scores.

The students in the lottery differ both academically and demographically from the students who were included in the regression discontinuity study. Not only do the lottery students have higher test scores than students at the eligibility cutoff, but their test scores exceed those of the average G&T student in the district. Lottery participants

No Measurable Benefit (Figure 2)

Participating in gifted programs had no impact on test scores among those barely eligible.



Note: The points show the authors' estimate of the impact of G&T program participation in each subject. The bars around each point show the range of possible effects that can be ruled out with 90 percent certainty. For example, one can be 90 percent certain that the effect of G&T participation on math scores is less than 0.04 standard deviations and greater than -0.11 standard deviations.

SOURCE: Author's calculations

are also less likely to be on subsidized lunch, and less likely to be minority.

Of the 542 lottery participants, only 440 students, including 331 winners (84 percent) and 109 losers (74 percent), remain in LUSD by 7th grade. Fortunately, the observed characteristics of lottery winners and losers who remain in the district continue to be very similar. Even so, when analyzing the data we control for students' demographic characteristics and prior achievement, and use weights designed to make the final sample comparable in terms of its observed characteristics to the set of students that initially applied for the lottery.

One disadvantage of this second study is that the lottery losers have a range of alternative experiences and most participate in standard G&T programs, so the comparison group's educational experience is less clear than it was in the regression discontinuity analysis. Nonetheless, our data confirm that students admitted to the G&T magnet schools with lotteries

seem to have experienced large improvements in their educational environment. Winning the lottery increased the average achievement of students' classroom peers by as much as a full standard deviation in some subjects. And in contrast to the G&T program as a whole, students admitted by lottery to G&T magnet program were assigned to more effective teachers.

Turning to student outcomes, however, our results provide little evidence that attending a G&T magnet program leads to improvements in student achievement (see Figure 3). The one exception is science test scores, for which we estimate a positive effect of 0.28 standard deviations. Due to the relatively small sample sizes, all of the effects are imprecisely estimated and do not allow us to definitively rule out reasonably large positive effects. Even so, the estimated effects for math, reading, and social studies are negative, and the estimated effect for language is effectively zero.

Discussion

It is difficult to understand why we find little evidence that G&T programs positively affect achievement. A common concern with studies of high-achieving students is that the available achievement measures may not be well suited to discern improvements for this group. This would be particularly worrisome if we were using a state accountability exam targeted toward low-achieving students, but it is less of an issue with the Stanford Achievement Tests. Indeed, we found little evidence of students performing near the maximum levels on these tests in either the regression discontinuity or lottery samples. Although it is possible that the additional course material taught in G&T classes is poorly aligned with topics covered in the achievement test, research documenting the benefits of being placed with higher-ability peers suggests that we should see improvements, even if that were the case.

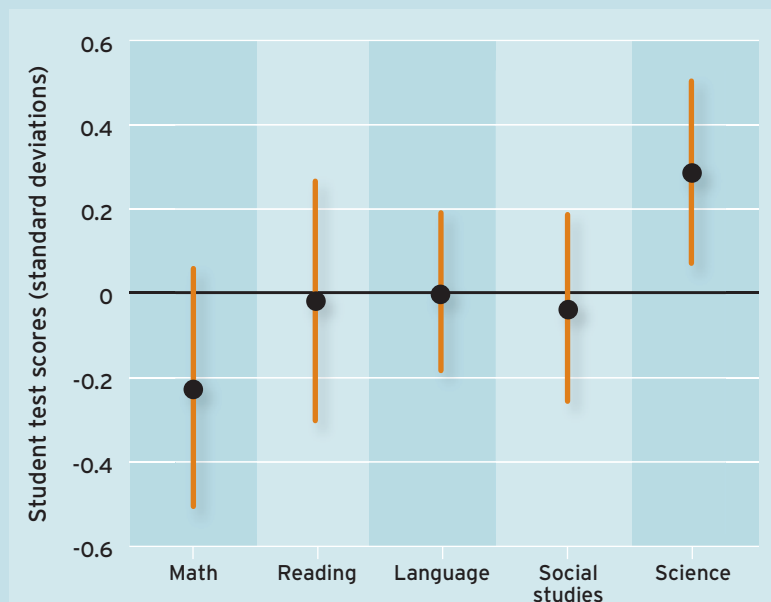
The effect of being placed in a higher-ability classroom may not necessarily be positive, however, especially for a marginal G&T student. In particular, the drop in ranking relative to one's peers may have a negative effect: a marginal G&T student is likely to go from being near the top of the regular class to being near the bottom of the G&T class. Even students in the middle of the G&T distribution are likely to experience a loss of ranking in the magnet G&T schools as compared to their neighborhood schools. It may be that students are demoralized by the drop in their relative rankings or that teachers provide more resources to students at the top of the class.

Substantial evidence from educational psychology indicates that students who are placed in higher-achieving groups can suffer psychological harm. A commonly used measure is a student's "self-concept," how a student perceives her abilities relative to an objective measure such as achievement. A 1995 study by Herbert Marsh and colleagues compared G&T students to observably similar students in mixed G&T and non-G&T classes and found that G&T students show declines in their math and reading self-concept. More recent research has documented lower self-concept and greater test anxiety among gifted students in ability-segregated classrooms.

Although we do not have direct evidence on student confidence, we can make use of student course grades and rank within the class to probe for evidence consistent with this kind of effect. We evaluate the impact of G&T program enrollment in the regression discontinuity study and of attending a G&T magnet in the lottery analysis. In both cases, we find clear reductions in student grades. For the regression discontinuity sample, grades fall by a statistically significant 4 points out of 100 (3 points changes a grade from a B+ to a B,

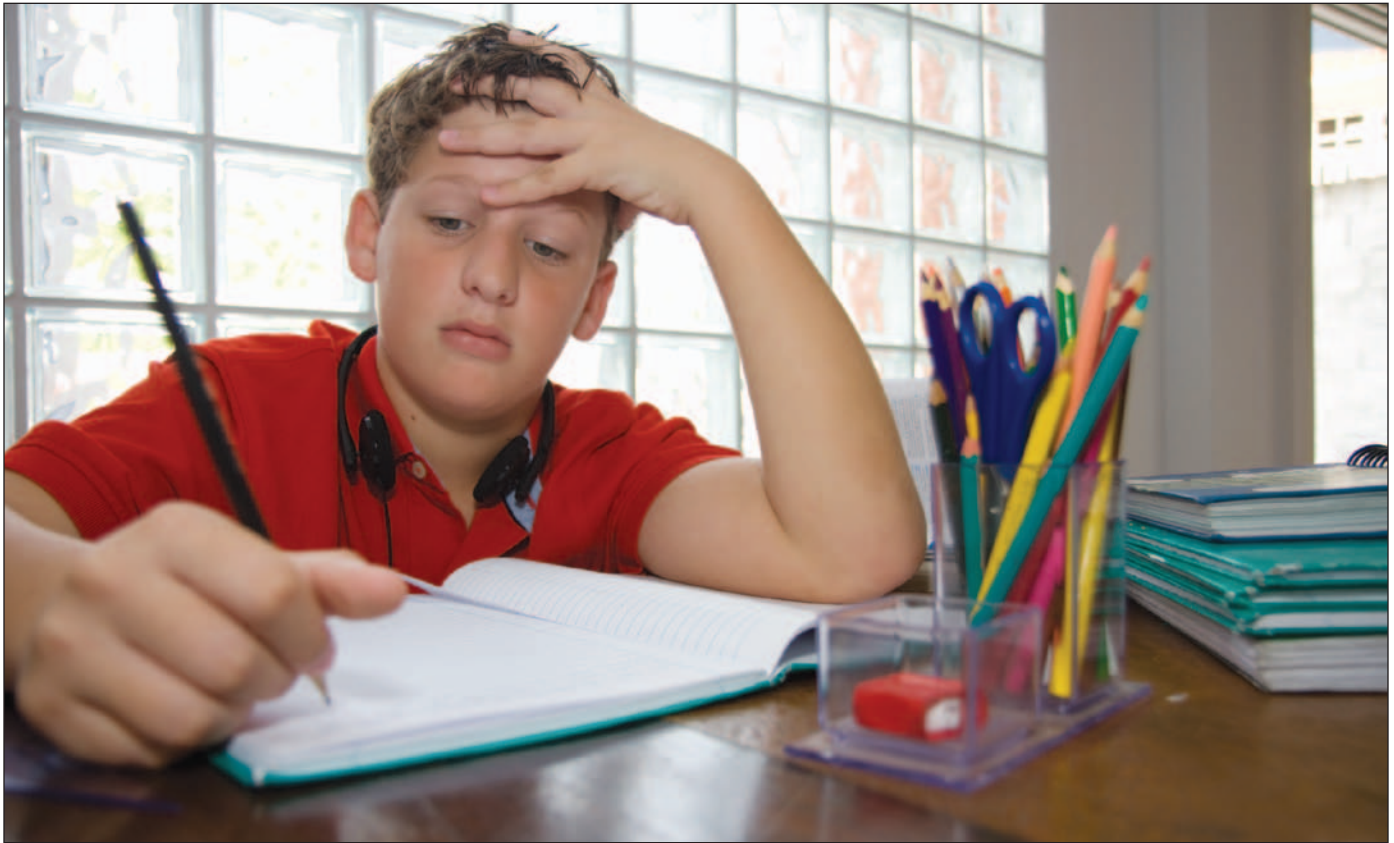
Stay In Your Neighborhood School? (Figure 3)

Attending a magnet gifted program increased science test scores but had no impact on scores in other subjects.



The points show the authors' estimate of the impact of G&T magnet program participation in each subject. The bars around each point show the range of possible effects that can be ruled out with 90 percent certainty. For example, one can be 90 percent certain that the effect of G&T magnet program on math scores is less than 0.06 standard deviations and greater than -0.50 standard deviations.

SOURCE: Author's calculations



for example) in math and by 2 to 3 points in other subjects, although these effects are not statistically significant for 7th grade. For the lottery analysis, the grade reductions are even more dramatic, with drops of 7 points in math, 8 in science, and 4 in social studies.

It is also useful to consider how students' rankings within their peer groups differ by treatment status, as this provides a direct measure of how a student may perceive his position in the overall distribution of student ability. We assume that students mostly compare themselves to their schoolmates who take the same courses in the same grade. Thus, we rank students within each school, grade, and course by their final course grades and then convert these rankings to percentiles. The rankings based on 7th-grade courses exhibit notable drops when students cross the G&T eligibility threshold. Controlling for race, gender, economic disadvantage, LEP (Limited English Proficiency), and prior gifted status, marginal G&T students have a relative rank in 7th grade that is 13 to 21 percentiles lower than similar students who were not admitted. Attending a premier G&T magnet in 7th grade generates a nearly 30 percentile ranking drop in all four of the courses examined.

In short, the necessary conditions are clearly met for a drop in relative ranking to play a role in offsetting the

expected positive impact of more rigorous courses, more effective teachers, and higher-achieving peers. The possibility that G&T students are subject to such a mechanism suggests potential constraints on the benefits of programs that provide more similar peers and an increase in traditional education inputs.

One should not conclude from the lack of achievement results, however, that the G&T programs should be scuttled. Our analysis occurs in a district with a large number of relatively high-quality magnet programs, and thus the alternatives to the G&T programs may be strong. There may also be benefits that we are not able to capture, such as impacts on SAT scores, graduation rates, and college attendance. Further, our study examines a G&T program in one district. Certainly, districts vary in the approaches they take to educating gifted students, so it may be that similar studies of programs in other districts would yield different results. Nonetheless, this study does raise questions about the efficacy of G&T programs and the traditional model of ability-segregated classrooms.

Sa Bui is a doctoral candidate in economics at the University of Houston, where Steven Craig is professor of economics and Scott Imberman is assistant professor of economics.