The Persistence of Early Scientific Interests: Evidence from the Science Talent Search

Ian H. Calaway
DePaul University

August 31, 2025

Abstract

This paper explores whether children demonstrating exceptional ability in science continue to pursue science into adulthood and whether their specific scientific interests persist through this transition. Using 80 years of data from the Science Talent Search (STS), a prestigious national science competition for high school students, I identify student characteristics associated with long-term scientific achievement. I find that students who pursue high school research projects in math, physics, astronomy, or computer science are significantly more likely to receive a National Science Foundation Graduate Research Fellowship (NSF GRF) than those whose projects focus on other disciplines, such as the life sciences. Moreover, these students tend to remain in the same scientific domains into adulthood: conditional on receiving an NSF GRF, they are most likely to receive it in the same field as their STS project. These findings suggest that the scientific interests of exceptionally talented students are quite "sticky."

1 Introduction

On December 28, 1942, President Franklin D. Roosevelt authorized the formation of the Manhattan Project with the goal of weaponizing nuclear energy. Although the Manhattan Project employed as many as 130,000 people at its peak, its success was largely dependent on the scientific breakthroughs of a handful of exceptionally talented scientists (Hewlett and Anderson, 1962; U.S. Department of Energy, 2000). The most indispensable among these include J. Robert Oppenheimer, Enrico Fermi, Glenn Seaborg, Ernest Lawrence, Edward Teller, Richard Feynman, Niels Bohr, and Leo Szilard. In the absence of any one of these individual scientists, the project may not have succeeded.

Given the key role star scientists play in innovation, its worth asking: what factors influence the production of such individuals? In previous work, I discussed the importance of mentors in identifying and nurturing students with exceptional math ability. In the context of science, one example of this is Richard Feynman. In a 1966 Citation for Distinguished Service in the American Journal of Physics, Richard Feynman wrote:

> *Abram Bader is cited for his superb teaching of an exceptional student. Usually, the truly outstanding stduent is left to go his own way. But when Bader found Rich Feynman (who this year shared the Nobel Prize in physics) in his class he gave the boy great challenges, good advice, and fascinating new information about physics. He excused him from paying attention in class but challenged him to use the time to learn advance calculus from a book from his personal library. He fascinated him by explaining to him the principal of least action, a central point of almost all of Feynman's work. Finally the student's love and admiration for his teacher resulted in the boy's wanting to become a teacher and scientist. Bader gave good counsel.*

From this passage it is clear that as a high school student, Feynman possessed scientific potential and a strong interest in physics (Feynman, 1966). From childhood to adulthood, Feynman exhibited *scientific persistence* and his specific interest in physics was *sticky*.

In this paper, I identify factors associated with scientific persistence in students who demonstrate exceptional science ability during high school. To do this, I leverage eighty years worth of novel data from a highly prestigious science competition, the Science Talent Search, covering more than 24,000 semifinalists across the US. I show that semifinalists whose projects in fields like mathematics, physics, astronomy, or computer science are significantly more likely to win a National Science Foundation Graduate Research Fellowship (NSF GRF) than peers whose projects focused on other areas, such as the life sciences. Moreover, I find that students tend to persist in the specific scientific domains they chose in high school: conditional on receiving an NSF GRF, they are most likely to win it in the same general field as their STS project. This serves as evidence that these exceptional students interests are sticky; like Feynman's early interests, they persist into adulthood.

2 Data

2.1 STS Data

> *The Science Talent Search was conceived and materialized in the belief that a major need for America today is the discovery and development of scientific ability among boys and girls now in high school. Real ability for creative research and engineering is rare. Although twenty Westinghouse scholarships are distributed among the Washington trip winners, the Science Talent Search is more than a scholarship contest. It is designed as a major step toward making available potential scientific talent to important tasks in war and peace.*
>
> — Watson Davis (1943)
> Director of the Science Service
> Founder of the Science Clubs of America

I constructed a novel dataset on exceptional science students by collecting and digitizing lists of approximately 300 semifinalists published annually by the Science Talent Search from 1942 to 2024. These lists, published in booklets until 1996, were distributed to colleges and universities across the US to help semifinalists secure admission and scholarships.[1] From 1997 onward, the lists were published on the organization's website.

These lists contain rich data on individual semifinalists. For each semifinalist, I observe first name, middle name or initial, last name, age, sex (until 1996), home address (until 1996), high school, city, and state. I use data from the Social Security Administration to impute sex based on students' first names for years after 1996. I leverage similar data from the 2000 and 2010 US Census to impute race based on students' last names.

Starting in 1958, the data include the title of each student's science essay or project. These titles provide unique insights into the specific scientific interests of these students at the end of high school. For example, Roger Y. Tsien, winner of the 2008 Nobel Prize in Chemistry, demonstrated his interest in chemistry at just 16 years old; the title of his Science Talent Search project was "Bridge Isomerism in Transition-Metal Thiocyanate Complexes."

To leverage the information implicit in these projects, I use a large language model (LLM) to sort these titles into nine general fields including Chemistry, Computer and Information Science; Engineering, Geosciences, Life Sciences, Mathematics, Physics and Astronomy, Psychology, and Social Sciences. These general fields were chosen based on fields historically recognized by the National Science Foundation Graduate Research Fellowship Program (NSF GRFP). In addition to

[1] For many years the collection of semifinalists was referred to as the *Honors Group*. The introduction of the 1981 booklet states, "The achieving of Honors in the Science Talent Search is an evidence of ability and interest in science and engineering which should be taken into consideration by colleges and universities in admissions and granting of scholarships. For this reason this Honors List is distributed to all universities and colleges of the United States. Each of those winning honors is being furnished with multiple certifications intended to accompany applications for colleges admissions and scholarships, but a listing in this pamphlet should be accepted as an equivalent certificate."

predicting project field, I also use the LLM to rate each project's technical difficulty, originality, and narrowness based on title.

2.2 NSF GRFP Data

I collected data on promising science graduate students by compiling records for the National Science Foundation Graduate Research Fellowship Program (NSF GRFP). The NSF GRFP has supported the graduate education of over seventy-five thousand US science students since 1952. These fellowships are highly competitive, with on average only 1,030 new fellowships awarded each year to graduate students across nine broad fields of science: Chemistry, Computer and Information Science, Engineering, Geosciences, Life Sciences, Mathematics, Physics and Astronomy, Psychology, and Social Sciences. The public NSF GRFP records list the first names, middle initials, last names, and field of graduate studies for all of these fellows.

I use these data to construct a conservative measure of scientific persistence. I consider this a measure of persistence because STS semifinalists who go on to receive NSF Graduate Research Fellowships are observably continuing their education in science at the graduate level. This is a conservative measure because NSF fellowships are highly selective; thus, the subset of STS semifinalists who receive them likely represents only a small portion of those who go on to attend graduate school in science. These data provide a unique opportunity to compare students' early scientific interests, as captured in the STS data, to their later fields of study as reflected in the NSF fellowship records.

2.3 Patents and Journal Publications

I plan to construct additional measures of scientific persistence for STS semifinalists by linking them to patent and publication records from Digital Science's Dimensions database. These data include over 100 million academic publications and more than 150 million patents, with detailed metadata on authors and inventors. Compared to the NSF GRFP data, these records should capture a broader share of STS semifinalists who continue in science, including those who pursue innovative careers outside academia. In addition to persistence, I will use these data to build measures of scientific productivity through publication and citation counts.

2.4 US Census Records

I plan to collect detailed information on the family backgrounds of the nearly 8,000 STS semifinalists born before 1950 by matching them to their 1940 and 1950 US Census records. Because I observe semifinalists' full names, ages, and hometowns, linking them to census records should be relatively straightforward. These census data will provide rich insights into their childhood environments, including parental education, occupation, income, and sibling composition. Given prior research emphasizing the influence of childhood environment on later scientific achievement, these data will

be valuable for controlling unobserved differences in students' propensities to pursue science. They will also be useful for benchmarking my estimates against existing literature.

3 The Composition of STS Semifinalists over Time

In this section, I use 80 years of student data from the Science Talent Search to highlight key patterns among high-achieving science students. Most notably, I document a large increase in the share of female students achieving semifinalist status, a dramatic rise in the share of Asian semifinalists during the 1980s, and gradual shifts in the types of projects students pursue.

Although these descriptive findings are drawn from a single academic competition, the long-standing prestige and national scope of STS suggest that these trends may reflect broader changes in the pipeline of high-achieving students. The long-standing prestige and national scope of STS suggest that these trends may reflect broader changes in the pipeline of high-achieving students. Although comparable long-term data for similarly prestigious programs like the National Merit Scholarship are unavailable, the STS provides a rare longitudinal lens on the shifting demographics and scientific interests of young talent in the US.

3.1 Trends in the Demographics of Semifinalists

In Figure 2, I show the trend in the share of male and female semifinalists over time. The female share rises steadily beginning in the 1960s, reaching gender parity for the first time in 2000. During this period of growth, there is a sharp 15pp decline between 1973 and 1974, coinciding with a change in the semifinalist selection policy. Prior to 1974, male and female students effectively competed separately, with the shares of semifinalists by gender set equal to the shares of competition entrants by gender. Consequently, the increase in the female share of semifinalists from 1942 to 1973 directly reflects an increase in the share of female entrants during the period.

I show the dramatic rise in the share of Asian semifinalists, from 10% in 1980 to over 60% by 2022, in Figure 3. The figure shows the share grew rapidly in two distinct periods. The first took place from 1980 to 1990 when the share of Asian semifinalists increased from 10% to 40%, followed by a share hovering between 30% and 40% through 2007. This pattern may reflect a delayed effect of the Immigration and Nationality Act of 1965, which significantly expanded immigration from Asia. The fifteen-year lag aligns with the children of immigrants reaching high school age. The second growth period begins in 2008, with the share climbing steadily and surpassing 60% by 2022. This later rise may similarly reflect a delayed effect of the Chinese Student Protection Act of 1992.

3.2 Trends in the Scientific Interests of Semifinalists

I show how the scientific interests of semifinalists have evolved from 1958 to 2025 in Figure 5. During this period shares of semifinalists with projects in chemistry, engineering, mathematics, and physics & astronomy have decreased, while the shares with projects in computer science, the life sciences, psychology, and the social sciences have increased. Throughout the period, projects in the

life sciences are by far the most common. Overall, the trends demonstrate that each cohort of semifinalists is diverse in terms of scientific interests. It remains unclear, however, how representative these shares are of the wider pool of competition participants.

In addition to changes in field composition, projects also appear to have become increasingly specific over time. This is apparent when reading project titles in chronological order, but providing direct empirical evidence of the trend is challenging. In Figure 6, I document one such measure by tracking the average length of project titles over time. From 1950 to 2020, the average length has roughly doubled. In Appendix .2, I present complementary evidence based on large language model (LLM) assessments of project difficulty, originality, and narrowness. I exclude these results from the main text due to the opaque nature of the scoring methodology, though they suggest similar patterns of increasing specificity.

4 Evidence on Persistence in Science

4.1 Summary of Semifinalist Sample with Persistence Measure

Students in the STS sample covering cohorts from 1958 to 2024 are demographically diverse and display a variety scientific interests in high school (Table 1). The sample is 36% female and 24% Asian with share for both of these demographics being high for later cohorts. The sample is also geographically diverse with representation in the Midwest, Northeast, South, and West US Census regions.

With respect to fields of STS projects, the students' high school science interests are also varied. The shares exceeding 5% of the sample for chemistry, computer science, engineering, life sciences, mathematics, physics & astronomy, and psychology. These interests are well represented throughout the sample (Figure 5), though life sciences is dominate with a 42% share during the period. Of the 18,703 semifinalists in the sample, 1,096 become NSF Graduate Research Fellows.

While the NSF Graduate Research Fellowship (GRF) rate conservative measure of science persistence, it masks important variation across scientific fields. To provide a more nuanced picture, Figure 7 presents a heatmap that illustrates how the scientific interests students express in high school, as captured by the field of their STS project, relate to the scientific field in which they ultimately receive an NSF fellowship.

Each cell in the heatmap indicates the number of semifinalists per 1,000 in a given STS field who go on to receive an NSF fellowship in a specific field. For example, among 1,000 students who conducted STS projects in the life sciences, on average 40 received NSF fellowships. Of those, about 40% of these fellowships were in the life sciences. Notably, students whose STS project was in mathematics have the highest overall fellowship rate, with 8.3 percent becoming fellows.

The figure reveals key patterns in the flows of students across scientific disciplines. The diagonal of the heatmap shows a tendency for students to remain in the same field. Conditional on becoming fellows, students are most likely to be awarded the fellowship in the same field as their STS project. This pattern holds true across all fields. The figure also highlights cross-field flows. For example,

Figure 1: Trends in the Demographics of STS Semifinalists

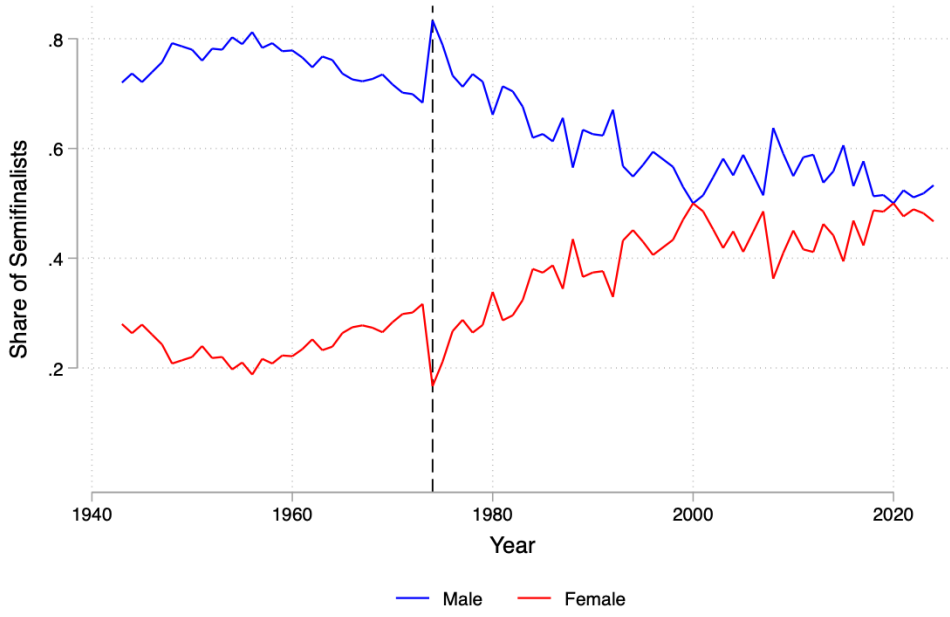


Figure 2: Trends in Semifinalist Gender

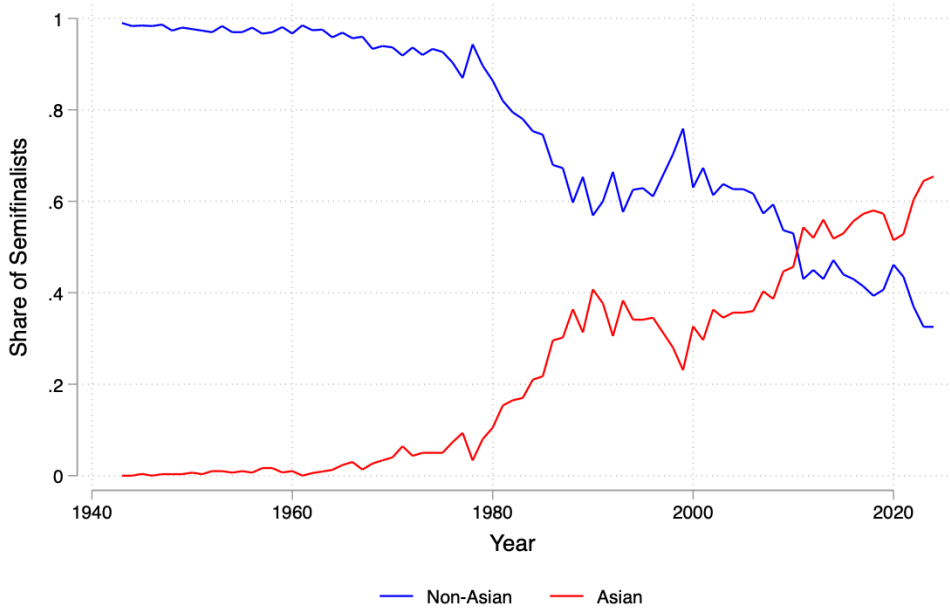


Figure 3: Trends in Semifinalist Race

Notes: Figure 2 shows the trend in the gender composition of STS semifinalists over time. The female share steadily increased starting in the 1960s, reaching gender parity for the first time in 2000. The dashed line at 1974 marks a policy change: prior to this year, the share of female semifinalists was set equal to the share of female competition entrants. Figure 3 displays the trend in the racial composition of semifinalists, highlighting a dramatic rise in the share of Asian semifinalists beginning around 1980.

Figure 4: Trends in the Projects of STS Semifinalists

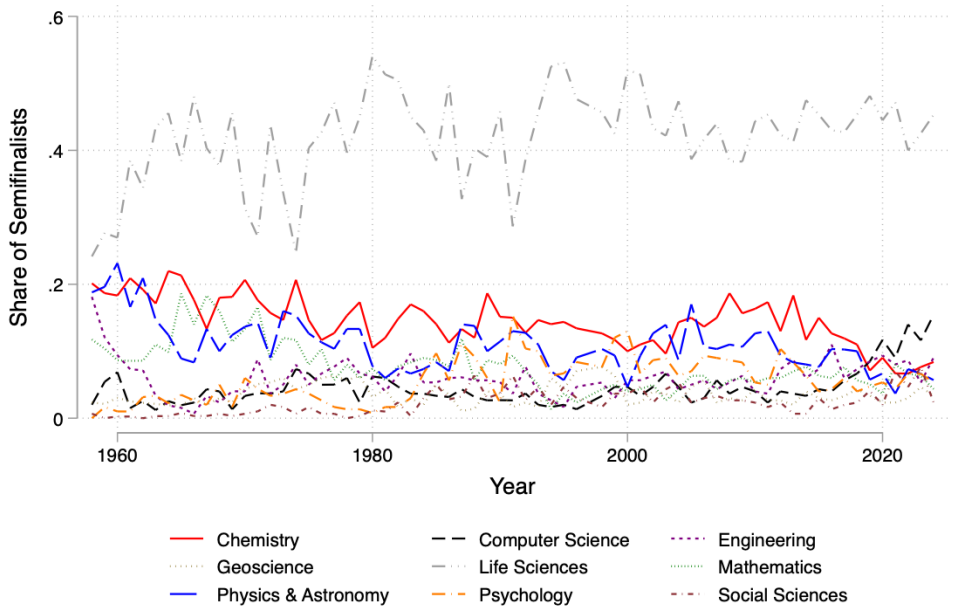


Figure 5: Scientific Interests of STS Semifinalists

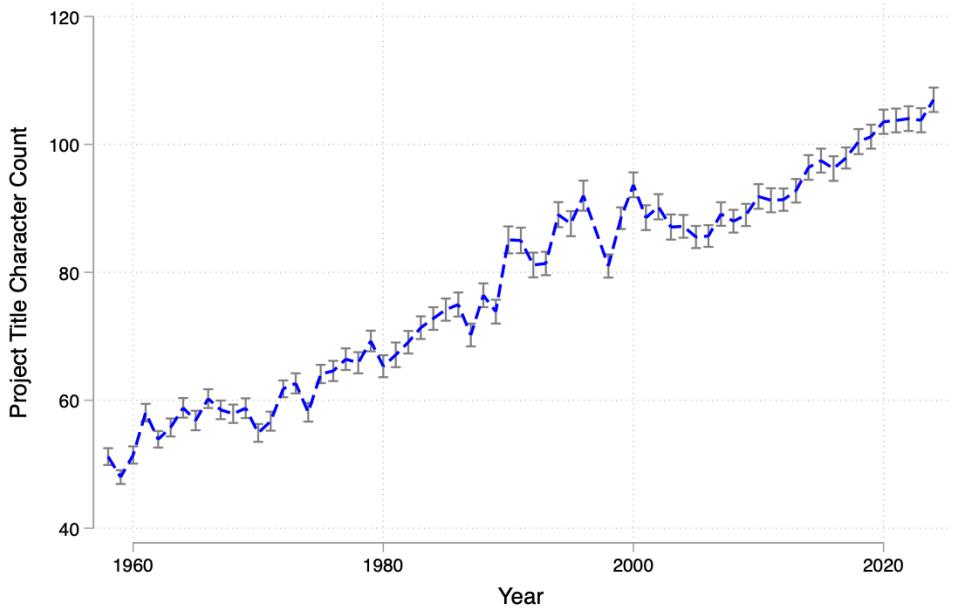


Figure 6: Average Project Title Length over Time

Notes: Figure 5 shows how the scientific fields of STS semifinalist projects have evolved from 1958 to 2025. Over this period, the shares of projects in chemistry, engineering, mathematics, and physics & astronomy declined, while the shares in computer science, the life sciences, psychology, and the social sciences increased. Life science projects remain the most common throughout. Figure 6 shows the average length of project titles from 1950 to 2020, which roughly doubles over time. This trend offers suggestive evidence that student research projects have become increasingly specific.

students whose STS projects were in the life sciences are relatively likely to remain in that field, whereas those from the geosciences are more dispersed across NSF disciplines. In the other direction, the column for geosciences NSF awardees suggests it attracts relatively few students with early interests in other fields. In contrast, engineering draws fellows from a broader range of early interests.

4.2 Estimating Persistence in Science

I test whether these persistence rates vary by these students' high school science interests, gender, and race. I do this by regressing an indicator for if the student received an NSF fellowship (y_i) on indicators for field (f) of project ($project_i^{(f)}$), student gender ($Female_i$), race ($Asian_i$), and covariates (X_{ist}) as well as school σ_s and year (τ_t) fixed effects. The specification is:

$$y_i = \alpha + \sum_{f \in F} \beta_0^{(f)} project_i^{(f)} + \beta_1 Female_i + \beta_2 Asian_i + X_{ist}\lambda + \tau_t + \sigma_s + \epsilon_{ist} \quad (1)$$

The models (Equation 1) are estimated using OLS for the three measures of persistence. The estimated coefficients of interest are displayed in Figure 8 and shown in Table 2; the reference group corresponds to a non-Asian, male semifinalist with a project in chemistry.

The results reveal meaningful demographic differences in long-term scientific persistence of students, even among this group of exceptional students. Female semifinalists are 1.7pp less likely than their male counterparts to become NSF Graduate Research Fellows. This is a substantial gap given the overall fellowship rate of just 6.4%. This gap is consistent with existing literature showing a leaky pipeline in STEM fields, where female students are more likely to switch out of science paths during post-secondary school.

The estimates also reveal racial gaps; Asian semifinalists are also less likely to win NSF fellowships, with a gap of 0.9pp relative to non-Asian peers. This corresponds to a 14% lower rate relative to non-Asian semifinalists. Given the rapid growth in Asian student representation among the STS semifinalists, this gap is surprising. It is unclear whether this stems from differences in occupational preferences, differential access to graduate programs, or other unobserved factors.

Differences by field of initial scientific interest, identified through STS project titles, are even more pronounced. Relative to semifinalists whose projects were in chemistry, those with mathematics STS projects are 7.9 percentage points more likely to become NSF fellows. In contrast, students whose projects were in the life sciences, psychology, or social sciences are significantly less likely to earn fellowships, with gaps ranging from 2.1 to 3.7 percentage points. Further work needs to be done to determine whether this actually reflects differences in student persistence or simply results from NSF funding priorities.

4.3 The Closing Gender Gap in Persistence

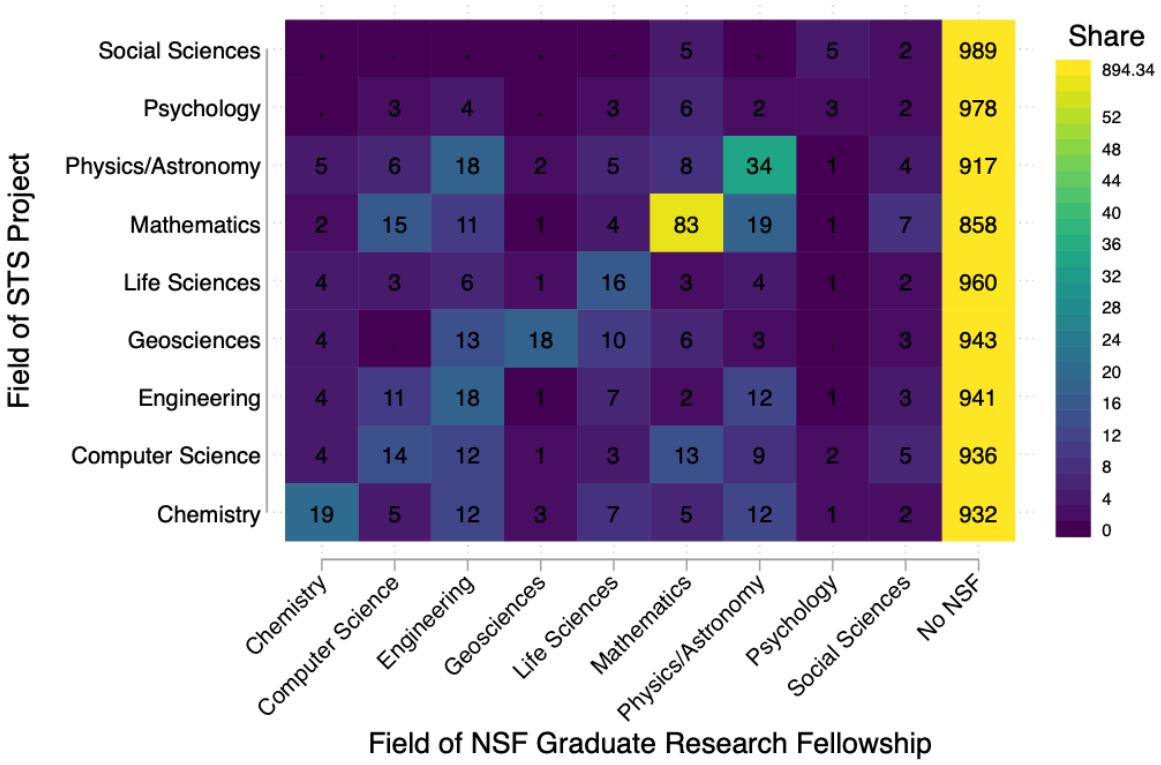
Given the massive increase in female labor force participation during the sample period, its worth testing whether female scientific persistence has changed over time. To do this, I modify the

Table 1: Summary Statistics for STS Student Sample (1958-2024)

	mean	sd
STS Cohort	1989.62	19.87
Female	0.36	0.48
Asian	0.24	0.43
Census Region		
Midwest	0.15	0.36
Northeast	0.47	0.50
South	0.24	0.43
West	0.14	0.35
STS Project Field		
Chemistry	0.15	0.36
Computer Science	0.05	0.22
Engineering	0.07	0.25
Geoscience	0.04	0.19
Life Sciences	0.42	0.49
Mathematics	0.08	0.27
Physics/Astronomy	0.12	0.32
Psychology	0.06	0.23
Social Science	0.02	0.15
Science Persistence		
NSF GRF	0.06	0.25
Observations	18,703	

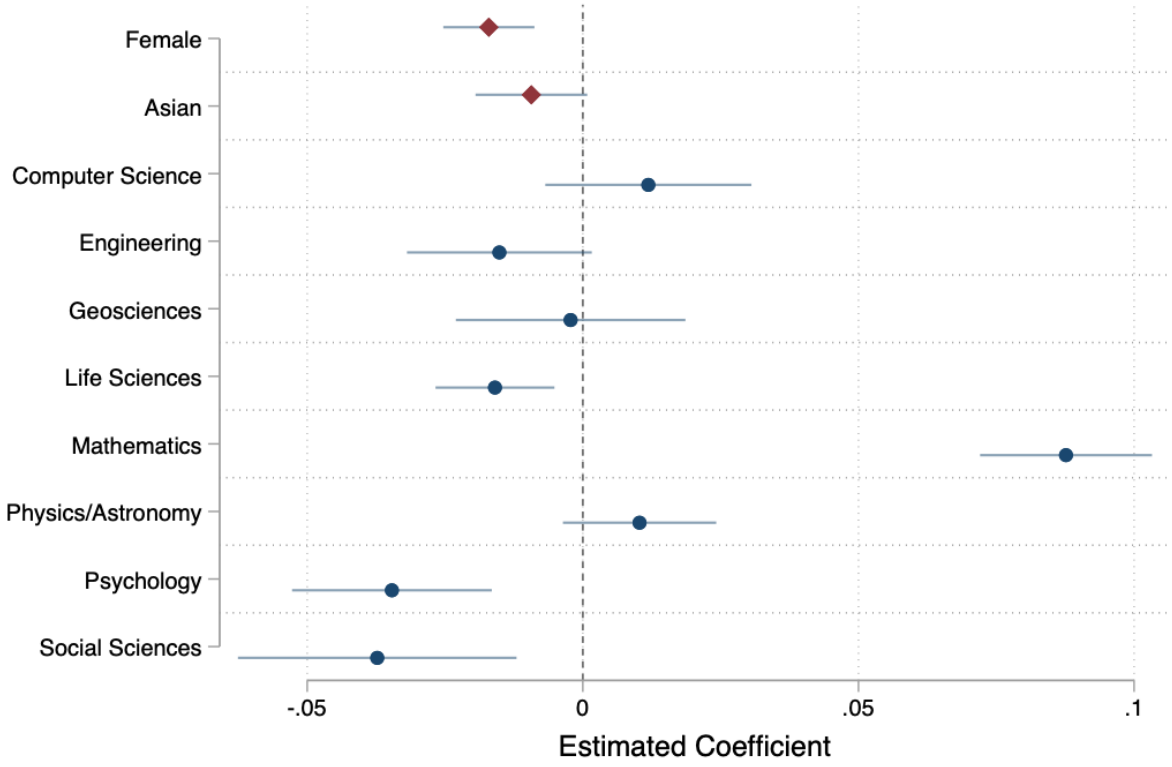
Notes: This table reports summary statistics for the 18,703 semifinalists in the Science Talent Search sample from 1958 to 2024. The sample is 36% female and 24% Asian. Students are geographically diverse, with representation across all four US Census regions. Their scientific interests in high school are varied, with more than 5% of students selecting chemistry, computer science, engineering, life sciences, mathematics, physics/astronomy, or psychology.

Figure 7: Field Transitions from STS Projects to NSF Fellowships



Notes: This figure shows how high school science interests, as measured by students' STS project field, relate to the scientific discipline in which they later receive an NSF Graduate Research Fellowship. Each cell shows the number of fellows per 1,000 semifinalists in the corresponding STS field. The diagonal reflects field persistence; off-diagonal entries capture movement across fields.

Figure 8: Semifinalist Predictors of NSF Graduate Research Fellowships



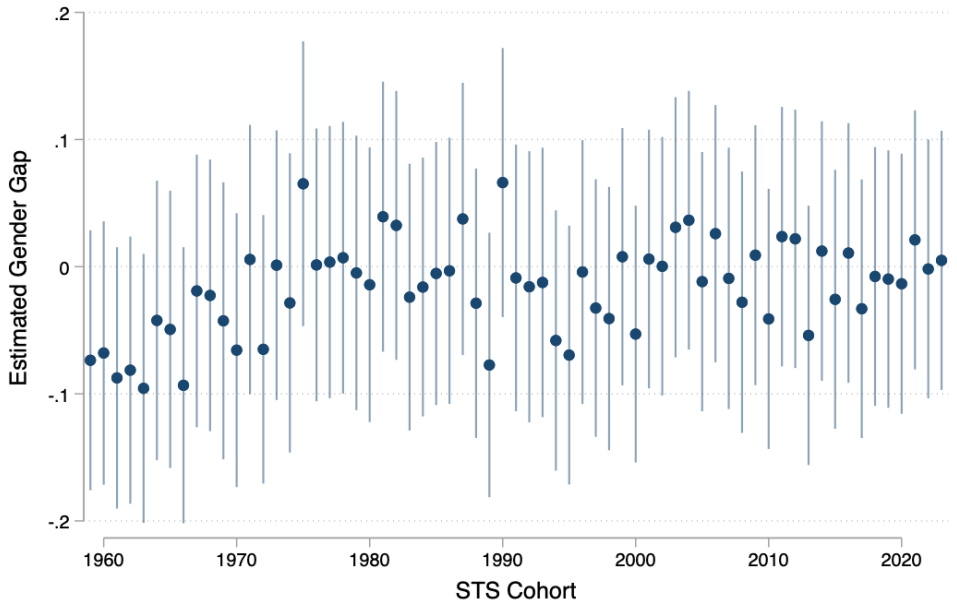
Notes: This figure shows estimated coefficients from OLS estimation of Equation 1, with 95% confidence intervals. The coefficients capture the association between gender, race, and high school science project field and the likelihood of receiving an NSF Graduate Research Fellowship (GRF). The reference group is a male, non-Asian student with a chemistry project. Students with projects in mathematics, physics, astronomy, or computer science are significantly more likely to receive GRFs than those with projects the life sciences, psychology, or social sciences. Female and Asian semifinalists are also significantly less likely to receive fellowships than their male and non-Asian peers.

previous specification by including interactions between gender ($Female_i$) and time (τ_t). The new specification is

$$y_i = \alpha + \sum_{t=1958}^{2024} \gamma^{(t)} Female_i \times \mathbf{1}[Cohort_i = t] + \sum_{f \in F} \beta_0^{(f)} project_i^{(f)} + \beta_1 Female_i + \beta_2 Asian_i + X_{ist}\lambda + \tau_t + \sigma_s + \epsilon_{ist} \tag{2}$$

As before, the model is estimated using OLS. The coefficients of interest in this case are those on the female and cohort interaction terms, the $\beta_0^{(f)}$ coefficients. When added to the coefficient on the female term, β_1, they capture the evolution of the gender gap in NSF fellowship recipience over time, while controlling differences in student race, hometown, and STS project field. This evolution in the gender gap is shown in Figure 9, and shows that the gender gap appears to close by the mid-1970s.

Figure 9: Closure of Gender Gap in Scientific Persistence



Notes: This figure shows the estimated coefficients on the interaction between gender and cohort from Equation 2, adjusted by the main female coefficient. It traces how the gender gap in NSF Graduate Research Fellowship receipt among exceptional students evolved over time. Estimates control for STS project field, race, hometown characteristics, and fixed effects for cohort and school. The results suggest that the gender gap present in the 1960s had largely closed by the mid-1970s.

5 Discussion

This paper's findings show that exceptional high school science students exhibit a strong tendency to persist in the scientific fields they initially pursue, as evidenced by their higher likelihood of receiving NSF Graduate Research Fellowships in the same general area as their original Science Talent Search projects. This persistence in early scientific interests underscores the importance of recognizing and supporting specific talents and passions during childhood. In future work, I will expand these findings by incorporating additional measures of persistence based on publication and patent data, which will capture a wider range of scientific careers beyond graduate study. I will also link these semifinalists to US Census records to observe their childhood environments, providing crucial controls for family background and socioeconomic factors. These extensions will deepen our understanding of the mechanisms that drive long-term scientific achievement and help benchmark the role of early scientific promise within broader social contexts.

References

Agarwal, Ruchir and Patrick Gaule, "Invisible geniuses: Could the knowledge frontier advance faster?," *American Economic Review: Insights*, 2020, *2* (4), 409–424.

Aghion, Philippe, Ufuk Akcigit, A. Hyytinen, and O. Toivanen, "The Social Origins of Inventors," *National Bureau of Economic Research*, 2017, (w24110).

Airoldi, Andrea and Paolo Moser, "Inequality in Science: Who Becomes a Star?," Technical Report, CEPR Discussion Paper No. 19572, Paris & London 2024.

Akcigit, Ufuk, John Grigsby, and Tom Nicholas, "The Rise of American Ingenuity: Innovation and Inventors of the Golden Age," *National Bureau of Economic Research*, 2017, (w23047).

Allensworth, Elaine M. and Kallie Clark, "High School GPAs and ACT Scores as Predictors of College Completion: Examining Assumptions About Consistency Across High Schools," *Educational Researcher*, 2020, *49* (3), 198–211.

Azoulay, Pierre, Joshua S. Graff Zivin, and Jialan Wang, "Superstar extinction," *The Quarterly Journal of Economics*, 2010, *125* (2), 549–589.

Beam, Margaret R., Chuansheng Chen, and Ellen Greenberger, "The Nature of Adolescents' Relationships with Their "Very Important" Nonparental Adults," *American Journal of Community Psychology*, 2002, *30* (2), 305–325.

Beilock, Sian Leah, "Reactivating the SAT/ACT requirement for Dartmouth undergraduate admissions," February 2024. Accessed May 31, 2025.

Bell, Alex and Neviana Petkova, "The Long-Term Impacts of Mentors: Evidence from Experimental and Administrative Data," Forthcoming. Working Paper.

_ , **Raj Chetty, Xavier Jaravel, Neviana Petkova, and John Van Reenen**, "Who becomes an inventor in America? The importance of exposure to innovation.," *The Quarterly Journal of Economics*, 2019, *134* (2), 647–713.

Bettinger, Eric P., Brent J. Evans, and Devin G. Pope, "Improving College Performance and Retention the Easy Way: Unpacking the ACT Exam," *American Economic Journal: Economic Policy*, 2013, *5* (2), 26–52.

_ , **Bridget Terry Long, Philip Oreopoulos, and Lisa Sanbonmatsu**, "The Role of Application Assistance and Information in College Decisions: Results from the H&R Block FAFSA Experiment," *The Quarterly Journal of Economics*, August 2012, *127* (3), 1205–1242.

Bilalić, M., K. Smallbone, P. McLeod, and F. Gobet, "Why are (the best) women so good at chess? Participation rates and gender differences in intellectual domains," *Proceedings of the Royal Society B: Biological Sciences*, 2009, *276* (1659), 1161–1165.

Bilen, Eren and Alexander Matros, "The Queen's Gambit: Explaining the superstar effect using evidence from chess," *Journal of Economic Behavior & Organization*, 2023, *215*, 307–324.

Bloom, Nicholas, Charles I. Jones, John Van Reenen, and Michael Webb, "Are Ideas Getting Harder to Find?," *American Economic Review*, 2020, *110* (4), 1104–1144.

Bressoud, David M., "The Strange Role of Calculus in the United States," *ZDM–Mathematics Education*, 2021, *53* (3), 521–533.

Buser, Thomas, Muriel Niederle, and Hessel Oosterbeek, "Gender, competitiveness, and career choices," *The Quarterly Journal of Economics*, 2014, *129* (3), 1409–1447.

California Institute of Technology, "Caltech Restores Standardized Test Requirement for Undergraduate Admission," April 2024. Accessed May 31, 2025.

Card, David and A. Abigail Payne, "High school choices and the gender gap in STEM," *Economic Inquiry*, 2021, *59* (1), 9–28.

_ **and Laura Giuliano**, "Universal screening increases the representation of low-income and minority students in gifted education," *Proceedings of the National Academy of Sciences*, 2016, *113* (48), 13678–13683.

Chetty, Raj, David J. Deming, and John N. Friedman, "Diversifying Society's Leaders? The Determinants and Causal Effects of Admission to Highly Selective Private Colleges," NBER Working Paper 31492, National Bureau of Economic Research June 2023. Accessed May 31, 2025.

_ , **John N. Friedman, and Jonah E. Rockoff**, "Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood," *American Economic Review*, 2014, *104* (9), 2633–2679.

_ , **Xavier Jaravel, Neviana Petkova, and John van Reenen**, "Do Tax Cuts Produce More Einsteins? The Impacts of Financial Incentives vs. Exposure to Innovation on the Supply of Inventors," *Journal of the European Economic Association*, 2019, *17* (3), 651–677.

Cimpian, Joseph R., Taek H. Kim, and Zachary T. McDermott, "Understanding persistent gender gaps in STEM," *Science*, 2020, *368* (6497), 1317–1319.

Colen, Liesbeth, Rene Belderbos, Stijn Kelchtermans, and Bart Leten, "Reaching for the stars: When does basic research collaboration between firms and academic star scientists benefit firm invention performance?," *Journal of Product Innovation Management*, 2022, *39* (2), 222–264.

College Board, "What is Digital SAT Adaptive Testing?," August 2023. Accessed May 31, 2025.

Committee on the American Mathematics Competitions, "Summary of Results and Awards," 1980–2023. Archive compiled by Ian H. Calaway.

Cook, Lisa D., Janet Gerson, and Jennifer Kuan, "Closing the innovation gap in pink and black," *Entrepreneurship and Innovation Policy and the Economy*, 2022, *1* (1), 43–66.

Corcoran, Sean P., William N. Evans, and Robert M. Schwab, "Women, the labor market, and the declining relative quality of teachers," *Journal of Policy Analysis and Management*, 2004, *23* (3), 449–470.

Dam, Levi Van, D. Smit, B. Wildschut, S. J. T. Branje, Jean E. Rhodes, Mark Assink, and Geert Jan JM Stams, "Does natural mentoring matter? A multilevel meta-analysis on the association between natural mentoring and youth outcomes," *American Journal of Community Psychology*, 2018, *62* (1-2), 203–220.

Dynarski, Susan, CJ Libassi, Katherine Michelmore, and Stephanie Owen, "Closing the Gap: The Effect of Reducing Complexity and Uncertainty in College Pricing on the Choices of Low-Income Students," *American Economic Review*, June 2021, *111* (6), 1721–56.

Ellison, Glenn and Ashley Swanson, "The Gender Gap in Secondary School Mathematics at High Achievement Levels: Evidence from the American Mathematics Competitions," *Journal of Economic Perspectives*, 2010, *24* (2), 109–128.

_ **and** _ , "Do Schools Matter for High Math Achievement? Evidence from the American Mathematics Competitions," *American Economic Review*, 2016, *106* (6), 1244–1277.

_ **and** _ , "Dynamics of the gender gap in high math achievement," *Journal of Human Resources*, 2023, *58* (5), 1679–1711.

Erickson, Lance D., Steve McDonald, and Jr. Elder Glen H., "Informal Mentors and Education: Complementary or Compensatory Resources?," *Sociology of Education*, October 2009, *82* (4), 344–367.

Falk, Armin, Fabian Kosse, and Pia Pinger, "Mentoring and schooling decisions: Causal evidence," Technical Report, IZA Discussion Papers 2020.

Fenoll, Ainoa Aparicio, Flavia Coda Moscarola, and Sarah Zaccagni, "Mathematics Camps: A Gift for Gifted Students?," *Journal of Economic Behavior & Organization*, 2021, *191*, 738–751.

Feynman, Richard, "1966 Citations for Distinguished Service," *American Journal of Physics*, 1966, *34*, 458–460.

Friedman, John N., Bruce Sacerdote, Douglas O. Staiger, and Michele Tine, "Standardized Test Scores and Academic Performance at Ivy Plus Colleges," *AEA Papers and Proceedings*, 2025, *115*, 676–681.

Fruiht, Veronica and Thomas Chan, "Naturally Occurring Mentorship in a National Sample of First-Generation College Goers: A Promising Portal for Academic and Developmental Success," *American Journal of Community Psychology*, 2018, *61* (3-4), 386–397.

Funk, Cary and Kim Parker, "Women and men in STEM often at odds over workplace equity: Perceived inequities are especially common among women in science, technology, engineering, and math jobs who work mostly with men," 2018.

Gowdy, Grace and Renée Spencer, "It's Who You Know That Matters: Identifying Which Type of Informal Mentor Is Most Likely to Promote Economic Mobility for Vulnerable Youth," *Journal of Primary Prevention*, June 2021, *42* (3), 217–238. Epub 2021 Mar 30.

Hagler, Matthew A. and Jean E. Rhodes, "The long-term impact of natural mentoring relationships: A counterfactual analysis," *American Journal of Community Psychology*, 2018, *62* (1-2), 175–188.

Hampole, Menaka V., "Financial Frictions and Human Capital Investments," Working Paper 2022.

Hanushek, Eric A. and Steven G. Rivkin, "Generalizations about using value-added measures of teacher quality," *American Economic Review*, 2010, *100* (2), 267–271.

_ , Marc Piopiunik, and Simon Wiederhold, "The value of smarter teachers: International evidence on teacher cognitive skills and student performance," *Journal of Human Resources*, 2019, *54* (4), 857–899.

Heckman, James J. and Stefano Mosso, "The Economics of Human Development and Social Mobility," *Annual Review of Economics*, August 2014, *6* (1), 689–733.

Herrera, Carla, Jean Baldwin Grossman, Tina J. Kauh, and Jennifer McMaken, "Mentoring in Schools: An Impact Study of Big Brothers Big Sisters School-Based Mentoring," *Child Development*, 2011, *82* (1), 346–361.

Hewlett, Richard G. and Oscar E. Anderson, *The New World, 1939–1946*, University Park, PA: Pennsylvania State University Press, 1962. Volume I of the official history of the U.S. Atomic Energy Commission.

Hoekstra, Hopi, "Return to Required Testing," April 2024. Accessed May 31, 2025.

Hoisl, Karin, Hans Christian Kongsted, and Monica Mariani, "Lost Marie Curies: Parental impact on the probability of becoming an inventor," *Management Science*, 2023, *69* (3), 1714–1738.

Hoxby, Caroline and Christopher Avery, "The Missing "One-Offs": The Hidden Supply of High-Achieving, Low-Income Students," *Brookings Papers on Economic Activity*, 2014, pp. 1–50.

Jaravel, Xavier, Neviana Petkova, and Alex Bell, "Team-Specific Capital and Innovation," *American Economic Review*, 2018, *108* (4-5), 1034–1073.

Jr., Robert E. Lucas, "On the Mechanics of Economic Development," *Journal of Monetary Economics*, 1988, *22* (1), 3–42.

Kraft, Matthew A., Alexander J. Bolves, and Noelle M. Hurd, "How Informal Mentoring by Teachers, Counselors, and Coaches Supports Students' Long-Run Academic Success," *Economics of Education Review*, 2023, *95*, 102411.

Laband, David N. and Bernard F. Lentz, "Like father, like son: Toward an economic theory of occupational following," *Southern Economic Journal*, 1983, pp. 474–493.

Lemann, Nicholas, *Higher Admissions: The Rise, Decline, and Return of Standardized Testing*, Princeton University Press, 2024.

Lindquist, Matthew J., Joeri Sol, and Mirjam Van Praag, "Why do entrepreneurial parents have entrepreneurial children?," *Journal of Labor Economics*, 2015, *33* (2), 269–296.

Lubinski, David and Camilla Persson Benbow, "Study of Mathematically Precocious Youth After 35 Years: Uncovering Antecedents for the Development of Math-Science Expertise," *Perspectives on Psychological Science*, 2006, *1* (4), 316–345.

_ , Camilla P. Benbow, and Harrison J. Kell, "Life Paths and Accomplishments of Mathematically Precocious Males and Females Four Decades Later," *Psychological Science*, 2014, *25* (12), 2217–2232.

McHale, John, Jason Harold, Jen-Chung Mei, Akhil Sasidharan, and Anil Yadav, "Stars as catalysts: an event-study analysis of the impact of star-scientist recruitment on local research performance in a small open economy," *Journal of Economic Geography*, 2023, *23* (2), 343–369.

Miranda-Chan, Thomas, Veronica Fruiht, Valeska Dubon, and Laura Wray-Lake, "The functions and longitudinal outcomes of adolescents' naturally occurring mentorships," *American Journal of Community Psychology*, 2016, *57* (1-2), 47–59.

Morgan, Allison C., Nicholas LaBerge, Daniel B. Larremore, Mirta Galesic, Jennie E. Brand, and Aaron Clauset, "Socioeconomic roots of academic faculty," *Nature Human Behaviour*, 2022, *6* (12), 1625–1633.

Murphy, Kevin M., Andrei Shleifer, and Robert W. Vishny, "The allocation of talent: Implications for growth," *The Quarterly Journal of Economics*, 1991, *106* (2), 503–530.

National Center for Education Statistics, "Integrated Postsecondary Education Data System (IPEDS)," 2020. U.S. Department of Education, Institute of Education Sciences.

Niederle, Muriel, "A gender agenda: A progress report on competitiveness," *American Economic Review*, 2017, *107* (5), 115–119.

_ **and Lise Vesterlund**, "Explaining the gender gap in math test scores: The role of competition," *Journal of Economic Perspectives*, 2010, *24* (2), 129–144.

Novosad, Paul, Sam Asher, Callum Farquharson, and Elion Iljazi, "Access to Opportunity in the Sciences: Evidence from the Nobel Laureates," 2024. Working Paper.

Quinlan, Jeremiah, "Test-Flexible Message to Counselors," February 2024. Accessed May 31, 2025.

Raposa, Elizabeth B., Jean Rhodes, Geert Jan J. M. Stams, Noel Card, Sophia Burton, Sarah Schwartz, Lori A. Y. Sykes, Stella S. Kanchewa, Janis Kupersmidt, and Saida Hussain, "The Effects of Youth Mentoring Programs: A Meta-analysis of Outcome Studies," *Journal of Youth and Adolescence*, 2019, *48* (3), 423–443.

_ **, Lance D. Erickson, Molly Hagler, Rachel Hamilton, Lillian T. Eby, Laura Burton, E. Alice Edwards, Megan Rainey, and Sarah E. O. Schwartz**, "How Economic Disadvantage Affects the Availability and Nature of Mentoring Relationships During the Transition to Adulthood," *American Journal of Community Psychology*, 2018, *61* (1-2), 191–203.

Rockoff, Jonah E., "The impact of individual teachers on student achievement: Evidence from panel data," *American Economic Review*, 2004, *94* (2), 247–252.

Rodríguez-Planas, Núria, "Longer-Term Impacts of Mentoring, Educational Services, and Learning Incentives: Evidence from a Randomized Trial in the United States," *American Economic Journal: Applied Economics*, 2012, *4* (4), 121–139.

Romer, Paul M., "Endogenous technological change," *Journal of Political Economy*, 1990, *98* (5), S71–S102.

Rosen, Sherwin, "The economics of superstars," *The American Economic Review*, 1981, *71* (5), 845–858.

Roy, Andrew Donald, "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 1951, *3* (2), 135–146.

Sacerdote, Bruce, Douglas O. Staiger, and Michele Tine, "How Test Optional Policies in College Admissions Disproportionately Harm High Achieving Applicants from Disadvantaged Backgrounds," Working Paper w33389, National Bureau of Economic Research 2025. NBER Working Paper No. 33389.

Schmill, Stu, "We are reinstating our SAT/ACT requirement for future admissions cycles," March 2022. Accessed May 31, 2025.

Stanford University, "Stanford will resume standardized test requirement for undergraduate admission," June 2024. Accessed May 31, 2025.

Surjadi, Milla and Sara Randazzo, "The Cheating Scandal Rocking the World of Elite High-School Math," *The Wall Street Journal*, August 15 2024. `https://www.wsj.com`.

Terman, Lewis Madison et al., *Genetic Studies of Genius: Mental and Physical Traits of a Thousand Gifted Children*, Vol. 1, Stanford University Press, 1926.

U.S. Department of Energy, "The Manhattan Project: An Interactive History," 2000. Accessed May 31st, 2025.

Wai, J., S.M. Anderson, and K. Perina, "The most successful and influential Americans come from a surprisingly narrow range of 'elite' educational backgrounds," *Humanities and Social Sciences Communications*, 2024, *11*, 1129.

Zimmerman, Marc A., Jeffrey B. Bingenheimer, and Diana E. Behrendt, "Natural mentoring," in David L. DuBois and Michael J. Karcher, eds., *Handbook of youth mentoring*, Sage Publications, 2005, pp. 143–157.

.1 Categorizing Science Projects using a Large Language Model

The following prompt was used with GPT-3.5 to extract information from project titles:
You are analyzing project titles from a prestigious U.S. high school research competition held from 1942 to 2024. Each title represents a high-level student research project across a wide range of scientific disciplines. Your task is to extract structured information from each project title.

1. **Fields**: Based on the title alone, choose the three most likely scientific fields the project belongs to, ordered from most to least likely. Select from this list of 9 fields (do not invent new ones): Chemistry; Computer and Information Science; Engineering; Geosciences; Life Sciences; Mathematics; Physics and Astronomy; Psychology; Social Sciences.

2. **Probability**: Estimate the probability (between 0 and 1) that the first field you selected is the same as the field the student officially selected in administrative records.

3. **Difficulty Score**: Based on the title alone, assign a score from 1 to 100 reflecting how technically difficult the project appears to be.

4. **Originality Score**: Based on the title alone, assign a score from 1 to 100 reflecting how original the project appears to be.

5. **Narrow Score**: Based on the title alone, assign a score from 1 to 100 reflecting how narrow the project appears to be in terms of content.

Respond only in the following structured format: Field: [field1, field2, field3] &&&& Probability: [field1_probability] &&&& Difficulty Score: [diffscore] &&&& Originality Score: [origscore] &&&& Narrow Score: [narrowscore]

Project title:

.2 Evidence on Persistence in Science Tables

Table 2: Coefficients for Persistence OLS Estimation

	(1)	(2)
Female	-0.0221***	-0.0170***
	(0.00387)	(0.00422)
Asian	-0.00537	-0.00933*
	(0.00473)	(0.00517)
Computer Science	0.0112	0.00159
	(0.00929)	(0.0101)
Engineering	-0.0143*	-0.0130
	(0.00834)	(0.00928)
Geosciences	-0.00281	-0.00735
	(0.0104)	(0.0116)
Life Sciences	-0.0177***	-0.0210***
	(0.00536)	(0.00588)
Mathematics	0.0881***	0.0789***
	(0.00776)	(0.00845)
Physics/Astronomy	0.0114*	0.0137*
	(0.00692)	(0.00762)
Psychology	-0.0336***	-0.0327***
	(0.00899)	(0.00966)
Social Sciences	-0.0372***	-0.0373***
	(0.0126)	(0.0132)
Time FE	STS Cohort	STS Cohort
Location FE	State	City
Mean	0.0642	0.0642
Observations	18,703	18,703
Adj. R2	0.0377	0.0529

Notes: This table reports the OLS estimated coefficients for Equation 1, which models the likelihood that a STS semifinalist becomes an NSF Graduate Research Fellow. The estimates show that even among these high-achieving science students, rates of becoming later receiving the fellowship vary by field of initial scientific interest, gender, and race. Female and Asian semifinalists are significantly less likely to later receive NSF fellowships than their male and non-Asian peers, respectively. Students with mathematics STS projects are the most likely to later receive these fellowships.