WHY COLLEGE ADMISSIONS SHOULD REMAIN TEST OPTIONAL/TEST FREE

(Despite What The New York Times Says)

Abstract

This FairTest report both deconstructs and responds to the New York Times "news analysis" on admissions testing policies AND presents a collection of studies, data, arguments, and resources in support of test optional/test free college admissions.

Harry Feder, Executive Director
harry@fairtest.org

Akil Bello, Senior Director
akil@fairtest.org

TABLE OF CONTENTS

THE ONTOLOGY OF COLLEGE ADMISSIONS EXAMS

The adoption of test optional and test free admissions policies by nearly 90% of four-year colleges is a clear financial threat to the College Board, ACT, and all others who feed off the admissions testing beast. Thus, it should come as no shock that those vested interests would engage in a public relations campaign to prop up the franchise. What is disturbing is that the American paper of record would so brazenly join the campaign. So before dissecting and refuting the advocacy of David Leonhardt's January 7, 2024 New York Times "news analysis" (and the various lower grade apologists for the tests that have and will continue to come in its wake), I think it's important to discuss a deeper ontological question: Why does the SAT (and ACT) exist? And why is it being hawked on the front page of the New York Times?

The premise of the SAT is to sort and funnel talent into the finishing grounds of the American "meritocracy." The argument for the test as a requirement is that it: a) accurately predicts "merit" (we'll call that the "validity claim"); and b) provides an avenue for talented members of the underclass to have access to the benefits of training to be a part of the leadership class of society (we'll call that the "equity claim"). I will for the time being skip over the eugenicist origins of standardized testing[1], its historic racism and classism and simply focus on these two current claims.

The validity claim is dependent on two ideas. First, that there is an objective idea of merit that warrants granting an individual a pathway to higher learning and status. This of course presupposes the notion that the granting of the privilege of "elite" higher learning is one to be rationed to only the "deserving", rather than one that should be broadly available to all who seek it. That is a deeply flawed premise. But if we are going to create a meritocratic hierarchy, we want to have the best and brightest on top.

What does the SAT (for this report we mostly focus on the SAT but the critiques and descriptions also largely apply to the ACT) actually test? Unclear. It certainly tests performance speed as it is timed and if you don't move fairly rapidly from question to question you won't do well. It tests knowledge of certain mathematical concepts that test makers presuppose were learned in middle and high school. (Thus, the curriculum, focus, and quality of your prior schooling does matter.) It tests your speed-reading ability and your ability to understand the literal meaning of what you read. It tests some

[1] See https://www.nea.org/nea-today/all-news-articles/racist-beginnings-standardized-testing

understanding of the English language and some analytical capacity. Keep in mind it tests all these things poorly because the multiple-choice framework and the way these questions are asked bring the element of test taking strategy (things like process of elimination and thinking like the test maker) into the equation. Does the SAT sort out our best and brightest? There are a lot more talents and abilities worthy of consideration in those we would want in "leadership" it doesn't test than it tests. I won't even attempt to list them all here.

Using the SAT as the gatekeeper for higher education turns out to test one thing above all else: existing station in life. Nobody contests that SAT scores correlate fabulously to family income and wealth and parental education levels. Those factors determine how and where you are educated before you apply to college. Mind training, intellectual and personal habits, and comfort with the underlying content that is developed over the course of years all funnel into greater likelihood of doing well on the SAT. There is also the additional factor of being able, at the moment of inflection, to hire high priced tutors to prepare for the exam. What the SAT, and standardized tests generally, seem to pick up better than anything is whether your origins lie in the winning side of the existing birth "meritocracy".

The SAT is an extraordinarily effective self-validation mechanism for an elitist "meritocracy" to continue to perpetuate itself. It is designed to maintain the existing class structure.

The Opportunity Insights study relied upon by the Times fully supports this claim. SAT scores give information about who does well at Ivy Plus institutions – which according to the study are grotesquely disproportionately dominated by students who come from the upper income brackets – and goes on to "leadership positions" in the meritocracy because it is mostly a test of whether you have been cultivated in that strata of American society in the first place. Incidentally, firms described as elite (and working for them as a measure of success) by the Opportunity Insights study are those that have lots of Ivy Plus graduates working there. Including the New York Times. So there is a circular logic that's hard to get around.

The SAT is an extraordinarily effective self-validation mechanism for an elitist "meritocracy" to continue to perpetuate itself. It is designed to maintain the existing class structure.

Given that the reporters and editors and publishers of the New York Times are products of that structure, it should shock no one that the newspaper would seek to defend a key gatekeeper.

THE PREDICTIVE POWER OF THE SAT ("THE VALIDITY CLAIM")

Most of the correlation of SAT/ACT scores to college performance are captured in other factors as well, notably high school grades. The actual supportable validity claim for the SAT, the one to which the College Board and ACT have historically mostly limited themselves, is that entrance exams provide some additional information about college freshman year grades–about .1 standard deviations–over and above high school grades. Incidentally, the College Board and ACT tend to limit themselves to claims about freshman grades for a reason; the grade correlation dissipates over the course of four years. Turns out once students who didn't have the benefit of the advantaged kind of earlier education get the hang of it, their grades fall in line with other predictive elements.

The question then becomes, as our Oregon State friend Jon Boeckenstedt puts it, "is the juice worth the squeeze?" Given the equity and applicant suppressive factors of **requiring** test submission (see our discussion below), 90% of four-year colleges have gone test optional or test free for students enrolling in 2024.

To extrapolate this data to make conclusions about standardized admission testing generally when these schools are operating in their own rarified world is statistical and ethical malpractice.

What is the "game changing news" that Leonhardt and the Times are touting that make failure to rely on test scores for admission a criminally negligent act? The study done by researchers at Opportunity Insights was of 12 highly selective schools: the so-called Ivy Plus. It concluded that within students of a given gender, family income level, race or ethnicity, and among students with the same grades in high school, students with the highest possible test score (i.e., SAT score of 1600 or ACT score of 36) achieve a first-year college GPA that is 0.43 points higher than students with an SAT score of 1200 or ACT score of 25.

The first obvious critique of this finding is that the Ivy Plus schools comprise .6% of the college students in America. If you broadened out to include other private institutions in the top 33 schools ranked by Opportunity Insights, you would be up to 1%. To extrapolate this data to make conclusions about standardized admission testing generally when these schools are operating in their own rarified world is statistical and

ethical malpractice. In fact, the report itself states: "We caution that our analysis applies only to Ivy-Plus applicants and the predictive power of test scores and GPAs may differ in other settings." Not that you would know it from the press clippings or PR machine, but Leonhardt himself buries the caveat towards the end of the piece: "The SAT debate really comes down to dozens of elite colleges, like Harvard, M.I.T, Williams, Carleton, U.C.L.A. and the University of Michigan."

Next question. Who are these students with 1200 SAT scores at Yale and Princeton? Well, Opportunity Insights also concluded that three things account for the concentration of the offspring of the top 1% and top income quintile at these schools: legacy admissions, "non-academic" factors (extracurriculars, etc.), and athletics. The Opportunity Insights data supports the conclusion that if the Ivy Plus schools are truly worried about low performing students, they would not put a thumb on the admissions scale for legacies and athletes. Furthermore, comparing scores 400 points apart on a 1600 scale is a very different exercise of predictive power than comparing students that are 30 points apart. The College Board quietly cautions that small differences in scores do not necessarily indicate differences in ability. In fact, on student reports they show that a student scoring a 1500 would likely score anywhere from 1460 to 1540 (that's the standard error of measurement) were they to take the test again. Given that at the 12 Ivy Plus colleges in the study Leonhardt cites 92% of enrolled students had a high school GPA of 3.75 to 4.0 (and if you got rid of athletes, donors' children and legacies that number would likely start higher) it's not unfair to say all enrollees are highly accomplished academically. For an admissions officer at Yale or Princeton, it may be that an SAT score provides some additional criteria (whether it's a justified one or not) to compare the 4.0s. But the marginal utility of test scores diminishes rapidly when the variations in GPA (together with course rigor) are greater or in a different category.

The claim the Times makes about standardized tests on the equity front, in addition to being misleading, is essentially smoke and mirrors hiding the real issue.

The one clear conclusion from the Opportunity Insights report is that if you got rid of legacies, donors' children, and recruited athletes from the Ivy Plus admitted students pool there would be greater socioeconomic equity in the student body at those schools. If Harvard really wanted to open doors for underrepresented students of merit, however measured, that's the ticket. The claim the Times makes about standardized tests on the equity front, in addition to being misleading, is essentially smoke and mirrors hiding the real issue. The mistake was recently replicated in the announcement and subsequent coverage of Dartmouth's change in testing policy.

A careful analysis of the Opportunity Insights report yields the realization that its claim of test score utility is vastly overstated. Professor of Public Policy and Governance at the University of Washington Jake Vigdor notes that while the Times calls the predictive power of SAT scores on college grades "strong", 80% of the variance in college grades are among students with similar test scores. Furthermore, the study asks the wrong question. Asking whether used alone SAT scores are a better predictor of college grades is a red herring; that's not the way admissions work. The real question is "compared to a prediction that ignores SAT scores, how much better is a prediction that uses them?" Additionally, admissions offices aren't limited to just grades and scores. Turns out, using the Opportunity Insights data, if you include factors like gender, who's a legacy, athlete, first generation applicant, underrepresented minority, etc., the predictive power of the SAT gets cut by more than a third, while grades maintain their predictive power. It's where a student went to high school (with those attending better resourced "elite" schools that have the best metrics) that actually has the most predictive power for first year grades. Not surprising since exposure to advanced curriculum and pedagogy matters.

Professor Vigdor used the Times article as an example for students in his quantitative and statistical analysis class of a news story in which data is exploited to support a foregone conclusion. Ya.

The pillorying of the University of California (UC) by the Times for going test free "despite its own data showing the predictive value of tests" is another example of study misuse. Leonhardt goes to great lengths to explain UC's refusal to engage with him for the article, implying they have something to hide. If you're UC and you know how the piece is going to turn out, why bother? Perhaps Leonhardt has actually read the indictments of the predictive value study he references and his anticipated silence on the part of UC is his preferred result.

The study relied upon by the Times to make its claim about the predictive value of the tests compared to grades in the UC system was

The real question is compared to a prediction that ignores SAT scores, how much better is a prediction that uses them.

Turns out, using the Opportunity Insights data, if you include factors like gender, who's a legacy, athlete, first gen applicant, underrepresented minority, etc., the predictive power of the SAT gets cut by more than a third, while grades maintain their predictive power.

just flat out wrong because of fundamental flaws in methodology. Saul Geiser, the former director of admissions research for the UC system, explained that the claim that SAT and ACT scores are superior to high school grades in predicting how students perform at UC was spurious. The study suffered from a classic methodological error: *omitted variable bias*. In this case in the zeal to prove the testing-college grade correlation, the UC "Task Force on Standardized Testing" forgot to account for the fact that the test predicts differently among different groups. Turns out compared to high-school grades, SAT/ACT scores are much more strongly correlated with student demographics like family income, parental education, and race/ethnicity. As a result, when researchers omit student demographics in their prediction models, the predictive value of the tests is artificially inflated. Berkeley economist Jesse Rothstein explained how this omitted variable problem inflates the predictive validity of SAT scores on freshman college grades by about 20%. When these student demographics are included in the model, the findings are reversed: High-school grades in college- preparatory courses are actually the stronger predictor of UC student outcomes. Whoops.

In fact, data and studies, none of which are mentioned in the Times, have been pretty consistent in concluding that high school grades are a better predictor of college performance. This is especially true if you consider the 99.4% of college students outside of the Ivy Plus schools. A study that followed over 55,000 graduates of Chicago public schools by University of Chicago researchers published in 2020 found that GPA was a much better predictor than ACT scores of both freshman year grades and, probably more importantly, college graduation rates. Simple logic can explain the outcome. According to the lead author of the study Elaine Allensworth: "GPAs measure a very wide variety of skills and behaviors that are needed for success in college, where students will encounter widely varying content and expectations. In contrast, standardized tests measure only a small set of the skills that students need to succeed in college, and students can prepare for these tests in narrow ways that may not translate into better preparation to succeed in college."

[S]tandardized tests measure only a small set of the skills that students need to succeed in college, and students can prepare for these tests in narrow ways that may not translate into better preparation to succeed in college.

A 2020 study done by the University of Wisconsin-Milwaukee found that: "Measures of high school GPA were the only unique predictors of college GPA. Both cumulative and senior year GPA were strong predictors of college GPA." A 2019 study by UC Davis researchers found that GPA is a stronger predictor than SAT scores of freshman year college grades. A NACAC sponsored study led by William Hiss, dean of admissions at Bates College, released in 2014 looking at the academic records of

123,000 students at 33 test optional schools (including Wake Forest, Bowdoin, Smith and Holy Cross), concluded that high school GPA is the strongest predictor of whether a student will fare well in college and ultimately graduate and that students who have strong grade point averages in high school are likely to do well in college *even if their standardized test scores are poor*.

But wait, there's more! A 2022 University of Tennessee report stated that the "ACT only adds predictive value in the top few HSGPA deciles, which is unhelpful in admissions decisions." The 2022 University of Iowa report showed that 39% of students with an ACT score between 15 – 17 (which is below the ACT average score of 20) and a HSGPA of 3.0 graduated at a higher rate than students with a comparable HSGPA and an ACT score of 33+. Further, the report concluded the "likelihood of graduating in four years was fairly consistent based on GPA, irrespective of the ACT score level." At Purdue, as a single factor, SAT scores are really weak in predicting graduation. If you added the other factors to the equation, SATs might even have negative correlation.

And for the definitive examination of college completion rates (arguably the most important result metric) we have the 2009 study undertaken by former Princeton president (for those who demand Ivy Plus pedigree) William Bowen and his colleagues, *Crossing the Finish Line* (Princeton Univ. Press), which compared the predictive validity of high-school GPA vs. SAT/ACT scores in a massive sample of students at 54 U.S. public universities:

> High-school grades are a far better predictor of both four-year and six-year graduation rates than are SAT/ACT test scores. ... The consistency of the results is extraordinary. In all but one of these more than 50 public universities, high- school GPA remains a highly significant predictor of six-year graduation rates after taking account of the effects of test scores. ... Test scores, on the other hand, routinely fail to pass standard tests of statistical significance when included with high school GPA in regressions predicting graduation rates (pp. 113-115).

We pile on here only because we can. And opponents of reducing reliance on testing in admissions have just blatantly ignored studies that run contrary to their position.

Bottom line: Why rely on an instrument like the SAT that has its class and race bias baked into the measure when other factors are less susceptible to a misleading but optically powerful numerical value and accompanying veneer of objectivity? Colleges don't need to use test scores to predict who will succeed in or be able to take advantage of the benefits of college.

Interestingly, on an episode of This American Life in 2021, Jeremiah Quinlan, head of Undergraduate Admissions at Yale, explained that the test optional policy spawned at Yale by the pandemic proved to them that doing admissions without the SATs was manageable:

> I can say that it has not been as disruptive as we had thought it was going to be. We have found that if you just spend a little bit more time looking at the transcript, the essays, letters of recommendation, or even an interview, you can find evidence of academic preparation or curiosity or excitement or fit for Yale that can make us confident in our ability to admit the right type of students.

Mr. Quinlan has had a recent change of heart. Yale will now require submission of test scores under a "test flexible" policy that, in a partial effort to mitigate the application chilling effect of an SAT/ACT requirement, allows for AP or IB tests in lieu of the SAT/ACT. He conceded that the 1000 or so students admitted to Yale without scores are doing quite well at Yale, but now has cited the "diamonds in the rough" equity argument for the shift. Yale, however, will likely lose qualified socioeconomically disadvantaged applicants who instead of not submitting scores will just not apply when presented with average test scores for admitted students on the Yale website. The shift was made with limited data from its test optional policy of three admission cycles. (Although Yale admitted more African-American students during the last two years than it ever had previously). Editors at the Times (Leonhardt, Yale '94) certainly helped pave the way for the change.

VALID FOR WHOM AND FOR WHAT

The truth is, making a decision about whether an individual warrants admission to a particular college is nuanced and complex and beyond a single test score. We don't want to "waste" the scarce benefit of a certain type of education on those who are not likely to be able to take advantage of it. But above a certain rough threshold of college GPA that shows ability to handle the work, can we say that a student with a 3.8 is getting more out of college training than a student with a 3.5? That's slicing the bologna awfully thinly. I don't think anyone is convinced that the difference between an A- and a B+ tells you who "better deserves" or has "benefited more from" that education.

If standardized testing can be said to be a valid determinant of admission qualification, it may be so for a narrow slice of the education world that requires special skills or knowledge in order to actually master higher level disciplinary material. If someone can't draw, attending the Rhode Island School of Design (RISD) would be a waste of time. Juilliard and other high-level schools of music require applicants to

demonstrate their musical talent and capacity to become a professional by prescreening videos and auditions.

The Times makes much of MIT's decision to go back and require standardized test scores. But MIT is really no different from Juilliard or RISD, except its special requirement is some basic ability to handle high level mathematics. If students who get less than a 650 on the SAT math section haven't been able to handle the mathematical rigors of MIT and don't persist in school according to its internal data, then there is an argument to disqualify applicants who do not pass that threshold. But that's not how their policy has been spun by the media; MIT's case is not a general argument for required testing at non-specialized programs. It's an undergraduate unicorn.

Of course there are other ways besides standardized tests, with its attendant inaccuracies, biases and roadblocks to get at what MIT needs in its student body–just ask CalTech. In an effort to make their admissions more diverse and open opportunities for poor and minority applicants, as well as use what they feel are better measures of mathematical aptitude, they have dispensed with the SAT/ACT requirement and look for other evidence of mathematical aptitude. Mr. Leonhardt pointedly did not talk to CalTech's admissions department to see how they do it. If MIT's class of 2023 was its most diverse class ever with a claim of 15% Black and 16% Latino students[2] and 20% Pell eligible students, that's mostly an indictment of the way it has done admissions in prior years. Somehow they weren't finding many "diamonds in the rough" in the years prior to the pandemic when they required testing.

> Colleges have missions beyond having their graduates ascend to "leadership" or make oodles of cash at the "best" firms

Besides the claim that standardized tests predict college grades and thus college accomplishment, the Times and Opportunity Insights extend the claim to correlation with success in life. Alas, how the Times and Opportunity Insights define "success" in life is quite a narrow formulation. I would hope there is success in life beyond being a Supreme Court Justice, U.S. Senator or working at McKinsey or the Times. By those criteria I'm an abject failure. Thankfully, I've learned to live with that.

Leonhardt posits that we need to identify and educate students most likely to excel so they "can produce cutting-edge scientific research that will cure diseases and accelerate the world's transition to clean energy." Those students could be said to be

---

[2] Incidentally, when you add up the percentages of student racial background from the MIT website or the Times article you get to 109%. Some students seem to be checking multiple boxes and MIT is double counting. This makes the claim difficult to verify.

successful, although as discussed above that's not how Opportunity Insights defines success. And that's not really the majority of students the Ivy Plus colleges are churning out these days. This argument supposes that a standardized test, which encourages narrow pedantic thinking, is key to discovering creative problem solvers and out of the box thinkers. There is no single metric that will help an admissions officer discover the next inventor. Evidence actually suggests that the metrics and social climate of these schools is likely to squash the creative pursuits that would lead to cutting-edge research, great innovation, or the public spiritedness that would lead to helping others in favor of pursuing the personal wealth of banking, consulting and the corporate ladder. Seems a couple of the biggest tech innovators (if you count Microsoft and Facebook as innovative) had to drop out of Harvard to make it work. And you don't need an Ivy Plus degree to be a successful innovator. Jensen Huang, founder of Nvidia (the chip maker powering AI) graduated from Oregon State. There is also a long list of changemakers who have been denied admission to the Ivy Plus. Famously, Stanford rejected Malala Yousafzai, a straight A student who'd won a Nobel prize, because she hadn't taken the SAT. An Ivy Plus credential is hardly the overriding ticket to great social impact.

Colleges have missions beyond having their graduates ascend to "leadership" or make oodles of cash at the "best" firms (defined as those populated by Ivy Plus graduates). At least I hope so. Creating intellectually thoughtful, morally centered and learned citizens, and skilled workers might be a better goal. Making sure that the benefits of a college education are visited upon those who need that degree for the economic and social opportunity it provides and won't be OK without it might also be an important goal.

THE VENEER OF OBJECTIVITY PROBLEM

Standardized test scores provide the appearance of being an objective measure of something. After all, everyone takes a similar test under near identical conditions. They are a "common yardstick." The results are "norm referenced" meaning the scores for everyone will fit a bell curve, with half the test takers doing better than the other half. It's a ranking and sorting mechanism. Results also have the power of being represented by a

> To use anything other than a holistic process that takes into account the individual circumstances of each applicant in the *high stakes* college admissions process is essentially invalid.

number score; numbers give off the appearance of legitimacy in the result because you can compare one test taker's number to another. A 1400 is always a higher score than a 1200.

The tests purport to measure a common skill or ability. Exactly what that skill or ability entails is not clear; the claim is some form of "college readiness." But pretty much no psychometrician can actually articulate how or why it does that.

And just because standardized test scores *seem* to be objective does not make them so.

As Steve Sireci, Executive Director of the Center for Educational Assessment at UMass Amherst explains, in order for an assessment to be valid for an individual it *cannot actually be exactly the same for everyone*. We must distinguish what a score tells us about an individual versus some conclusion scores allow for the population of test takers as a whole. Sireci writes:

> [I]n educational testing, *students* are the most important part of the measurement process, not the measure itself, or the measurement scale. Contemporary psychometrics and educational research have clearly determined that overly rigid testing procedures can impede accurate measurement of students' proficiencies, and distort test score interpretations . . .
>
> [T]he United States has a richness of diversity with respect to language, history, and culture. How- ever, the large-scale educational measurement community did not emerge from this rich diversity of culture; rather, it emerged from the dominant culture—from those who were in power in the early 20th century. Those who were not in power were easily marginalized. Thus, the culture of educational testing in the United States today, grew out of the dominant culture of the times from the early-to-mid 20th century . . .
>
> [V]alid interpretation of students' test scores requires understanding the heterogeneity of the student population with respect to community resources, home resources, family structures, culture, language, communication norms, religious beliefs, educational experiences, and other factors. By understanding the different "funds of knowledge" (González, Moll, & Amanti, 2005) students bring to the testing situation, we can better standardize that situation to support, rather than prohibit, diversity. Standardization should not "wash out" student heterogeneity, it should embrace it.

This analysis leads to the conclusion that to use anything other than a holistic process that takes into account the individual circumstances of each applicant in the

high stakes college admissions process is essentially invalid.[3] And because of human, American, and institutional difference and the admissions test being the product of a particular culture and mode of thinking, it is not an objective measure. College admissions tests would have to be much more extensive, nuanced, and long-term assessments in order to come close to validly assessing student capabilities for something as broadly defined as success in college.

While far from perfect themselves, high school grades and teacher evaluations at least look at student performance over a long period of time subject to the scrutiny of multiple examiners and thus are more able to reflect and capture student nuance and difference in circumstance and talent. They at a minimum reflect levels of perseverance and executive function. Plainly that is more reflective of the capacity for "success" over time than a 2-3 hour exam.

THE DIAMONDS IN THE ROUGH ("THE EQUITY CLAIM")

The College Board and pro-testing crowd are smart enough to realize that given the history of standardized testing, and the stark disparate impact evident in the scoring data against socioeconomically disadvantaged and minority students, they must conjure up some kind of "opportunity" argument. Thus, the industry "equity claim" is that we need standardized tests to catch the poor and minority "diamonds in the rough." Leonhardt makes the broad claim that "test scores can be particularly helpful in identifying lower-income students and underrepresented minorities who will thrive." This is a very poor and flimsy argument for requiring standardized testing for college admissions. **The tests hurt the chances of far more poor and underrepresented students of talent than they help. Far more human**

[3] The professional standards governing the use of standardized tests established by the American Educational Research Association (AERA) indicate that tests should not be the sole determinant for high stakes education decisions. https://www.aera.net/About-AERA/AERA-Rules-Policies/Association-Policies/Position-Statement-on-High-Stakes-Testing. "Decisions that affect individual students' life chances or educational opportunities should not be made on the basis of test scores alone. Other relevant information should be taken into account to enhance the overall validity of such decisions. As a minimum assurance of fairness, when tests are used as part of making high-stakes decisions for individual students such as promotion to the next grade or high school graduation, students must be afforded multiple opportunities to pass the test. More importantly, when there is credible evidence that a test score may not adequately reflect a student's true proficiency, alternative acceptable means should be provided by which to demonstrate attainment of the tested standards."

potential is left on the shelf because of the tests than is uplifted by their identifying and sorting properties.[4]

Let's begin with Christina Paxson, President of Brown, who the Times quotes as saying that test scores are a better predictor of academic success at Brown than grades. Here are some other things she says that didn't make it into the Times. Paxson recognizes that high school grades are exceptional for the vast majority of Brown students but worries about grade inflation (we won't discuss this claim, but the origin of this claim comes from and supports the testing industry[5]). The rarified status of the Ivies is on display. On the equity front, here is her position:

> "there are clear drawbacks to requiring standardized tests. Simply put, students are less likely to apply to colleges that require test scores . . . Our biggest challenge will be ensuring that students we very much want push the "submit" button on applications. Requiring test scores could work against us by reducing the size and diversity of the applicant pool."

Suspending testing requirements, while not the magic bullet of educational opportunity, are good for equity. A study by Christopher Bennett of Vanderbilt University examined a diverse set of nearly 100 private institutions that adopted test-optional undergraduate admissions policies between 2005–2006 and 2015–2016 and found that test-optional policies were associated with a 3% to 4% increase in Pell Grant recipients, a 10% to 12% increase in first-time students from underrepresented racial/ethnic backgrounds, and a 6% to 8% increase in first-time enrollment of women. These patterns were generally similar for both the more selective and the less selective institutions examined. The William Hiss led NACAC sponsored study also concluded that optional testing policies help build broader access to higher education with non-submitters more likely to be first-generation-to-college students, minorities, Pell Grant recipients, women and students with learning differences. At Wake Forest University, non-submitters of test scores are twice as likely to be first-generation college students, Pell-eligible and/or domestic students of color; in other words, some of the most underrepresented and underserved students in higher education.

---

[4] Stories of actual students can help understand how this plays out in real life. Paul Tough's narrative of Daniela, a Mexican immigrant in Riverside, CA is one such compelling story.

[5] See self-interested research by the ACT-- https://www.act.org/content/dam/act/secured/documents/pdfs/Grade-Inflation-Continues-to-Grow-in-the-Past-Decade-Final-Accessible.pdf (for a critique see https://www.edsurge.com/news/2022-05-16-act-says-grade-inflation-is-a-serious-problem-it-s-probably-not) and the College Board, Hurwitz and Lee, *Grade Inflation and the Role of Standardized Testing*, in Measuring Success: Testing, Grades and College Admissions, (Johns Hopkins Press) (2018). See https://hechingerreport.org/proof-points-new-evidence-of-high-school-grade-inflation/

And those students who enter without test scores at Wake wind up doing just as well, or by some measures better, than students who submit scores. While after the first year of Wake Forest coursework achieved an average GPA 0.13 (out of 4.0) higher than test-optional students, the difference progressively shrinks to 0.03 by graduation. Even more impressive, a larger percentage of test optional students persist to graduation (90 percent) than their test-submitting counterparts (87 percent). The tests aren't needed to identify talented, successful students, and by removing the barrier, racial and socioeconomic equity is served.

Who are these "diamonds in the rough" whose scores outperform their high school GPA in statistically significant ways? If they were mostly poor minority students then perhaps the pro-test requirement crowd would have a real equity argument. But alas, and not surprisingly, they are not. Who has higher SAT scores and lower grades? Men, white and Asian students, wealthier students, and children of well-educated parents. Guess who tends to have higher grades and **lower scores**? Women (who get better grades at every level of education than men, by the way), poorer students, students from underrepresented ethnic groups, and students whose parents have less education.[6] This narrative plays smoothly into the origin story of the SAT, where those groups were thought not to be suited for higher education. No doubt there are true anecdotes of disadvantaged students being "found" by an SAT or ACT score. But the data is clear–far more are lost because of an SAT requirement. And "diamonds" who are "found" by test scores are more likely to be socioeconomically *advantaged* kids. That's a sizable contributor to why "elite" college campuses look the way they do as Opportunity Insights earlier research points out.

Policies that use GPA are much better at capturing socioeconomically disadvantaged and underrepresented minority students who could thrive in challenging academic undergraduate settings. Studies of so-called top grade percentage plans in both California and Texas prove this. Zach Bleemer, an economics professor at Princeton (so he must be OK) studied the efficacy of test-based "meritocracy" in college admissions by evaluating the impact of a grade-based "top percent" policy implemented by the University of California. The Eligibility in the Local Context (ELC) program provided large admission advantages to the top four percent of 2001- 2011 graduates from each California high school (thus minimizing test score impact considerably). Because top graduates from higher-performing high schools had little need for ELC eligibility to gain UC admission, 80 percent of barely eligible ELC participants were from the bottom half of California high schools by SAT. The ELC led over 10 percent of barely eligible applicants from low-opportunity high schools to enroll at selective UC

[6] The data supporting this can be found on the Higher Ed Data Stories blog here. Also see Akil Bello's analysis of discrepant scores data provided by the ACT and SAT here.

campuses instead of less selective public colleges and universities. Half of those participants came from lower-income families, and their average SAT scores were at the 14th percentile of their UC peers. Despite this seeming mismatch, ELC participants overperformed in their college grades; enrollment at the more selective UCs led participants to graduate earlier and earn higher late-20s wages by over $1,000 per percentage point change in their enrollment institution's graduation rate.

The students with discrepantly low SAT scores in comparison to their grades were from overwhelmingly more disadvantaged populations and lower test performing high schools. Their freshman year grades lagged behind their higher performing peers but outpaced their position predicted by SAT score. Despite their relatively poor initial academic preparation and performance relative to their more-advantaged peers, over the course of their time they caught up. ELC participants became about 0.8 percentage points more likely to earn a college degree within five years per 1 percentage point increase in the graduation rate of their enrollment institution (which rose by about 22 points overall). Bleemer concludes:

> These findings suggest that expanding selective university access to low-SAT high-GPA applicants – as by top percent policies, test-optional admissions (Belasco et al., 2015; Bennett, 2022), or holistic review (Bleemer, 2023) – could promote economic mobility without decreasing universities' average economic value-added to their enrolled students.

Providing access to high level college education to students with high grades but lower SAT scores gets you **success**–as measured by college completion and wages earned nearly a decade after graduation–for the students provided with the opportunity. Colleges get greater fairness in admissions by providing access to an ignored and marginalized pool of talent without taking the hit to their prestige they mistakenly fear.

A 2020 study published a year ago on the University of Texas 10% plan (TTP–where all students whose grades placed them in their school's top decile were guaranteed admission to the most-selective campuses) by Sandra Black (Columbia), Jeffrey Denning (BYU) and Jesse Rothstein (Berkeley) concluded that for highly ranked students at more disadvantaged high schools who gained access under the policy, college enrollment and graduation increased. Less highly ranked students at more advantaged high schools, who tended to lose access to the flagship campus shifted toward less-selective colleges under the policy, but did not see declines in overall college enrollment, graduation, or earnings. The policy thus benefited students targeted for admission without evidence of adverse effects on displaced students.

The Pulled In students (those from more racially diverse, higher needs high schools) who attended UT Austin as a result of TTP had graduation rates comparable to

the average UT Austin student, suggesting that despite their poorer high school preparation on average, these students were not mismatched. The Pushed Out students (those from more advantaged high schools who did not finish in the top 10% of their class and otherwise did not get into UT Austin) did not suffer from lower graduation rates or reduced wages. This suggests that the benefits of attending a more-selective public institution may be quite small for these students.

The policy based on grades was a driver of equity and opportunity for more disadvantaged and racially diverse students. *This automatic access was done without any consideration of standardized test scores.* Many Pulled In students would not have attended any college absent the policy. This shift substantially increased the share of students who earn bachelor's degrees, with no indication that these students suffered from attending more-selective colleges. In contrast, Pushed Out students lost access to UT Austin but offset this with higher enrollment rates at less-selective campuses, with no change in overall college enrollment. There was no evidence of negative effects on graduation rates or earnings for this group. The effectively test free admissions policy had real benefit for lower income and minority applicants without actually reducing opportunity for those who had greater advantage going in. Win win.

TEST-OPTIONAL MEANS "DIAMONDS IN THE ROUGH" CAN STILL BE FOUND

This idea is self-explanatory. A discrepant student (one who has a mismatch of grades and test scores) who has a lower GPA and gets a good score is welcome to submit it. Perhaps some of those students won't take the test unless they are made to, but the opportunity remains. And the data indicates requiring the test will turn away so many more otherwise qualified disadvantaged applicants. One can only conclude that the real motivators of the test requirement crowd are: a) keeping their *alma maters* "elite" (and all the coding that comes with that); b) enrollment management or c) economic self-interest.

THE EQUITY PROBLEM OF OTHER CRITERIA

The last gambit of the pro-testing crowd is to throw shade on other holistic criteria used by admissions offices as skewing the admitted student population towards the wealthy and advantaged. The Opportunity Insights study cites three factors in Ivy Plus schools admitting twice as many higher income students: athletic recruitment, legacy admissions, and higher non-academic ratings. The capacity of wealthier students to pad their resume with trips to plant trees in the Brazilian rainforest or attend the New York Times summer journalism program gives them an unfair advantage on the non-academic side of it.

Under the Opportunity Insights formulation, we would get two-thirds of the way home on the remedying the admissions bias in favor of higher income students by eliminating legacy preferences and athletic recruitment. The third factor, the non-academic ratings equity problem, could be remedied by admissions offices doing a better job in using an equity lens and being more critical of the activities submitted to produce this rating. How does one compare the after-school job at the 7-11 with the spring break in Costa Rica working on an organic farm? The calculation of the non-academic rating can be adjusted to equate cooking dinner for the family while mom is working a late shift with the Yo-Yo Ma cello lessons. The other part of the non-academic rating that skews wealthy is school counselor and teacher recommendations. Private school personnel have far smaller caseloads and more time for extensive individual attention to an individual student. As opposed to public high school teachers, private school teachers have more capacity to write lengthy well-crafted recommendation letters. These differences could be accounted for in calculating the non-academic score. Rather than revert to a metric that has its inequity baked into the instrument and its scoring (SAT/ACT–just look at the data), at least the non-academic ratings could be adjusted to reflect a different value system than currently animates them.

I would posit this to the proponents of testing–how come the equity and inclusion record on American "elite" campuses was so poor before the institution of test optional and test free admissions?

Can we invoke the ancient adage two wrongs don't make a right? Just because there is an admissions metric that produces an inequitable result doesn't mean you should keep another one that also excludes worthy socioeconomically disadvantaged candidates.

CONCLUSION

The vast majority of colleges in the United States do not require submission of a standardized test score for admission. They have come to the understanding that requiring tests operates as a barrier to otherwise qualified candidates from disadvantaged and underrepresented backgrounds in higher education, and that by relying on grades and other holistic considerations, they can select students for enrollment who will thrive in college and benefit greatly from advanced education. Test scores are unnecessary.

Now, armed with a study examining under 1% of the undergraduate population at the most selective schools in the country, where all enrollees other than many of the ones who have a special pass due to legacy, donor, or athletic status have extraordinarily strong credentials, the New York Times is effectively leading the charge for schools to break ranks and reestablish the SAT and ACT as gatekeepers. It has done so as an advocate for a policy, evidenced by its skewed analysis that fails to recognize the research and data that indicates there are real benefits for marginalized students to not require standardized tests, without compromising the academic mission and standard of the university. Doing so will cause many talented underrepresented students to simply not apply to these schools. Test optional and test free policies have uniformly increased applications from socioeconomically disadvantaged students. Given the close correlation of standardized tests to family wealth, reliance on that instrument for admissions will produce the same inequity in college admissions it has before colleges began abandoning the requirement.

I would posit this to the proponents of testing–how come the equity and inclusion record on American "elite" campuses was so poor before the institution of test optional and test free admission? How can you claim that this time will be different?

The data on relying on grades as the academic metric for admission has proven effective on the "success" front and superior in producing more socioeconomically diverse campuses. Why would schools (as Dartmouth just did) reverse course a mere two application cycles (impacted by COVID with students who have yet to graduate) into the policy? Because they were comfortable with the ranking and sorting of the SAT in the first place which for decades has successfully replicated student bodies of privilege that can be called "elite." It's all about reputation and standing.

The cold truth is this entire conversation focuses on the wrong thing. Testing is the tip of the iceberg. Colleges need to worry about the affordability of their education to the average high school graduate. They should not bestow a valued seat to legacies

nor place a thumb on the scale for a wealthy squash player. We certainly need improvements and greater equity in K-12 education. Selective colleges can and must do the deeper work of finding and encouraging potential and promise in students from all over the country and not just reserve slots from Exeter, Horace Mann and Harvard-Westlake. Requiring testing sends a signal to those students existing in worlds far removed from Ivy Plus corridors that undermines that task.

References

Research Reports

American Educational Research Association. (2000). Position Statement on High Stakes Testing. https://www.aera.net/About-AERA/AERA-Rules-Policies/Association-Policies/Position-Statement-on-High-Stakes-Testing

Allensworth, E. M., & Clark, K. (2020). High School GPAs and ACT Scores as Predictors of College Completion: Examining Assumptions About Consistency Across High Schools. Educational Researcher, 49(3), 198-211. https://doi.org/10.3102/0013189X20902110; https://www.aera.net/Newsroom/High-School-GPAs-and-ACT-Scores-as-Predictors-of-College-Completion-Examining-Assumptions-about-Consistency-across-High-Schools

Bennett, C. T. (2022). Untested Admissions: Examining Changes in Application Behaviors and Student Demographics Under Test-Optional Policies. American Educational Research Journal, 59(1), 180-216. https://doi.org/10.3102/00028312211003526 https://journals.sagepub.com/doi/10.3102/00028312211003526

Black, S., Denning, J., Rothstein, J. (2020). Winners and Losers? The Effect of Gaining and Losing Access to Selective Colleges on Education and Labor Market Outcomes. Working Paper 26821 https://journals.sagepub.com/doi/10.3102/00028312211003526 https://www.nber.org/digest/jun20/results-texass-experiment-increasing-college-diversity

Bleemer, Z. (2022). Affirmative Action, Mismatch, and Economic Mobility after California's Proposition 209. The Quarterly Journal of Economics, Oxford University Press, 137(1), 115-160.

Bleemer, Z. (2024). Top Percent Policies and the Return to Postsecondary Selectivity. https://zacharybleemer.com/wp-content/uploads/2020/10/ELC_Paper.pdf

Board of Regents, State of Iowa. (2022). Board of Regents Admissions Policy Changes. https://www.iowaregents.edu/media/cms/0122_ITEM_9__Admissions_Policy_Rec_4B94DAE5833BF.pdf

Bowen, W., Chingos, M., McPherson, M. (2009). Crossing the Finish Line: Completing College at America's Public Universities. Princeton Univ. Press.

Chetty, R., Deming, D., Friedman, J. (2023). Diversifying Society's Leaders? The Determinants and Causal Effects of Admission to Highly Selective Private Colleges.

Opportunity Insights. https://opportunityinsights.org/wp-content/uploads/2023/07/CollegeAdmissions_Paper.pdf

Galla, B. M., Shulman, E. P., Plummer, B. D., Gardner, M., Hutt, S. J., Goyer, J. P., Duckworth, A. L. (2019). Why high school grades are better predictors of on-time college graduation than are admissions test scores: The roles of self-regulation and cognitive ability. *American Educational Research Journal*, *56*(6), 2077-2115.

Geiser, S. (2020). SAT/ACT Scores, High School GPA, and the Problem of Omitted Variable Bias: Why the UC Taskforce's Findings are Spurious. Berkeley Center for Studies in Higher Education, Research & Occasional Paper Series: CSHE.1.2020. https://cshe.berkeley.edu/sites/default/files/publications/2.rops.cshe.1.2020.geisersatactommitted_variables.3.18.2020.pdf

Gilman, L., Jones, C., Davis, G.S. (2020). What Matters for College Success? The Relationships between GEAR UP Participation, High School GPA, AP Participation, and ACT Scores with College GPA. https://uwm.edu/sreed/wp-content/uploads/sites/502/2020/05/GU-Predictive-measures-of-postsecondary-success.pdf

Hiss, W., Franks, V. (2014). Defining Promise: Optional Standardized Testing Policies in American College and University Admissions. https://www.luminafoundation.org/files/resources/definingpromise.pdf

Hurwitz, M., Lee, J. (2018). "Grade Inflation and the Role of Standardized Testing," in Measuring Success: Testing, Grades and College Admissions, Johns Hopkins Press. 64-78.

Kurlaender, M., Cohen, K. (2019). Predicting college success: How do different high school assessments measure up? [Report]. Policy Analysis for California Education. https://edpolicyinca.org/publications/predicting-college-success-how-do-different-high-school-assessments-measure-2019

Rothstein, J., (2004). College Performance Predictions and the SAT. Journal of Econometrics, 121, 1–2, 297-317. https://www.sciencedirect.com/science/article/abs/pii/S0304407603002537?via%3Dihub

Sanchez, E., Moore, R. (2022). Grade Inflation Continues to Grow In the Past Decade. An ACT Report. https://www.act.org/content/dam/act/secured/documents/pdfs/Grade-Inflation-Continues-to-Grow-in-the-Past-Decade-Final-Accessible.pdf

Sireci, S.G. (2020). Standardization and *UNDERSTAND*ardization in Educational Assessment. Educational Measurement: Issues and Practice, 39: 100-105. https://doi.org/10.1111/emip.12377

The University of Tennessee Board of Trustees. (2022). Standardized Testing in Admissions Supplemental Information. https://trustees.tennessee.edu/wp-content/uploads/sites/3/2022/02/ERS-Committee-Supplemental-Information.pdf

Articles, Blogs, and Podcasts

https://akilbello.com/2020/10/16/dont-believe-the-hype/

Barshay, J. "Proof Points: New Evidence of High School Grade Inflation," The Hechinger Report, May 16, 2022. https://hechingerreport.org/proof-points-new-evidence-of-high-school-grade-inflation/.

Glass, I.. Tough, P. "The Campus Tour Has Been Cancelled," This American Life. Episode 734. March 19, 2021. https://www.thisamericanlife.org/734/transcript

https://www.highereddatastories.com/2019/03/looking-at-discrepant-scores.html

https://jonboeckenstedt.net/2021/06/07/the-status-quo-strikes-again/

Leonhardt, D. "The Misguided War on the SAT,: The New York Times, January 7, 2024. https://www.nytimes.com/2024/01/07/briefing/the-misguided-war-on-the-sat.html?searchResultPosition=1

McGuire, E. "Are You Considering Test Optional Admissions," Inside Higher Education, August 28, 2022. https://www.insidehighered.com/admissions/views/2022/08/29/test-optional-admissions-works-opinion

Noonoo, S. "ACT Says Grade Inflation Is a Serious Problem. It's Probably Not." EdSurge, May 26, 2022. https://www.edsurge.com/news/2022-05-16-act-says-grade-inflation-is-a-serious-problem-it-s-probably-not)

Paxson, C. "To Test or Not To Test: A Letter From the President," Brown Alumni Magazine, June-August 2023. https://www.brownalumnimagazine.com/articles/2023-06-20/to-test-or-not-to-test