

## Contextualizing the SAT: Experimental Evidence on College Admission Recommendations for Low-SES Applicants

**Michael N. Bastedo (corresponding author)**

University of Michigan  
Center for the Study of Higher and Postsecondary Education  
610 E. University Ave., Ann Arbor, MI 48109  
[bastedo@umich.edu](mailto:bastedo@umich.edu)

**Kristen M. Glasener**

[glasener@umich.edu](mailto:glasener@umich.edu)

**KC Deane**

[deanek@umich.edu](mailto:deanek@umich.edu)

**Nicholas A. Bowman**

[nick-bowman@uiowa.edu](mailto:nick-bowman@uiowa.edu)

### Abstract

Although it is well established that college entrance exams have become a key factor for admission to selective institutions, less is known about the influence of test scores in relation to other academic factors in the evaluation of a student's application file. This study conducts a randomized-controlled trial to determine whether providing students' test scores in context—how they perform relative to their school and neighborhood peers—increases the likelihood that admission officers ( $n = 321$ ) would recommend admitting low-socioeconomic status (SES) applicants. The study also examines how including a personal admission essay that conveys grit, or ability to persevere in the pursuit of long-term goals, influences admission decision making. Admission officers in the contextual condition were significantly more likely to accept both the low-SES and high-SES applicant than those without contextual information on test scores; however, they were not more likely to accept applicants who convey grit in their personal essays.

**Keywords:** college admissions, college access, low-income students, grit, randomized controlled trial, holistic review

**Citation:** Bastedo, M. N., Glasener, K. M., Deane, K. & Bowman, N. A. (2019). Contextualizing the SAT: Experimental evidence on college admission recommendations for low-SES applicants. *Educational Policy*. Advanced online publication. <https://doi.org/10.1177/0895904819874752>.

### Acknowledgements:

We are particularly thankful for the support and feedback of Jessica Howell, Michael Hurwitz, and Greg Perfetto.

### Funding

This paper was written with the generous research support of The College Board. KC Deane received training support from the Institute of Education Sciences, U.S. Department of Education (PR/Award R305B150012#). Any opinions, findings, and conclusions or recommendations

expressed in this material are those of the authors and do not necessarily reflect the views of The College Board or the Institute of Education Sciences.

### **Author Biographies**

Michael N. Bastedo is a professor of education and the director of the Center for the Study of Higher and Postsecondary Education at the University of Michigan. His scholarly interests are in higher education decision making, particularly college admissions, stratification, enrollment management, rankings, and governance.

Kristen M. Glasener is a doctoral candidate in higher education in the Center for the Study of Higher and Postsecondary Education at the University of Michigan. Her research focuses on selective college admissions policies and issues of college access and diversity.

KC Deane is a doctoral candidate in the Center for the Study of Higher and Postsecondary Education at the University of Michigan. Prior to arriving at Michigan, she worked as a program manager at the Aspen Institute's College Excellence Program and as a research associate at the American Enterprise Institute's Center on Higher Education Reform. She holds a BA in economics from Reed College and a master's in education policy from the University of Pennsylvania's School of Education.

Nicholas A. Bowman is a professor of Higher Education and Student Affairs and the director of the Center for Research on Undergraduate Education at the University of Iowa. His research uses a social psychological lens to explore salient issues in higher education, including student success, diversity, undergraduate admissions, college rankings, and research methodology.

Selective colleges and universities are learning how to grapple with two potentially competing trends in college admission. The first trend involves an increased reliance on standardized test scores as a central component used to determine whether to admit an applicant (Alon & Tienda, 2007). Standardized test scores not only help ease the burden of evaluating growing applicant pools (Lemann, 1999) but they are also frequently cited as an important predictor of students' future performance in college (Shaw, Kobrin, Patterson, & Mattern, 2012). When used without contextual factors influencing a student's performance, such as the academic opportunities available at their high schools, test scores may become a powerful sorting mechanism that disproportionately affects the chances of admission for applicants from low-socioeconomic status (SES) backgrounds.

Although reliance on standardized test scores has increased over time, a second trend in selective college admissions is the growing pressure to improve the socioeconomic diversity of institutions' student bodies. From the White House College Opportunity Summit in 2014 to recent efforts like the American Talent Initiative, top-tier colleges and universities are being asked to increase the number of low-SES students in their campuses. Part of this pressure is the result of modest improvements in low-SES student enrollment over the past several decades. Low-income students (those in the bottom quintile of income) remain severely underrepresented at selective institutions, constituting less than 5% of enrollment at the most selective institutions (Chetty, Friedman, Saez, Turner, & Yagan, 2017). Although low-income students have made strong gains in academic preparation since the 1970s, higher-income students have made even stronger gains on the same academic indicators (Alon & Tienda, 2007; Bastedo & Jaquette, 2011). As a result, low-SES students remain at a competitive disadvantage in the selective

admissions process compared with their higher-SES peers, and over-reliance on standardized test scores in admissions may exacerbate these inequalities.

Persistently low enrollment among low-SES students at selective institutions is not due to a lack of qualified candidates. A substantial number of low-SES high school graduates each year earn standardized test scores that are typical of highly selective colleges (Hoxby & Avery, 2012). Other students have scores that are somewhat lower but still impressive when compared with their school and neighborhood averages. Despite attending less-resourced schools and growing up without the economic advantages of their higher-SES peers, many socioeconomically disadvantaged students perform extremely well in school and are capable of being academically successful at selective institutions if offered admission (Bowen & Bok, 1998).

Within the larger applicant pool at selective institutions, students from low socioeconomic backgrounds are frequently compared with their higher-SES peers, who are more likely to live in neighborhoods composed primarily of high-SES households and attend well-funded schools with long-established college-going cultures (Frank & Cook, 1995; Reardon & Bischoff, 2011). Such students likely appear better-qualified for admission to selective colleges; however, if high-achieving, low-SES students were compared with the academic performance of their similarly situated peers, perhaps a greater number of low-SES students would be admitted to selective colleges and universities. One open question is whether admission officers would be more likely to admit low-SES students if they had more robust contextual information regarding their test score performance.

To enhance access for disadvantaged students, many selective institutions have implemented admission strategies such as holistic review or the assessment of non-cognitive factors. Although there is no agreed-upon definition of holistic admission, it generally refers to

admission officers' efforts to contextualize their review of a student's credentials, incorporating information about an applicant's family and school context that might influence their learning opportunities (College Board, 2002; Lucido, 2014). Evidence suggests that providing contextual information regarding students' family background and high school context increases the likelihood of admitting lower-SES applicants (Bastedo & Bowman, 2017). Admission officers may also use holistic review to gauge the presence of desirable non-cognitive factors, such as persistence or self-discipline (Hossler et al., 2019; Sternberg, 2010), and look for compelling narratives and storytelling through the essay (Stevens, 2007). However, it is unclear whether providing information on students' non-cognitive traits in the selective admission process influences the relative evaluation of lower- and higher-SES student applicants. Without such evidence, it is not clear if these policies help improve access to selective institutions for low-SES students.

Therefore, in this study, we explore whether providing contextual information related to standardized test scores and non-cognitive factors influences admission officers' decision making. Specifically, we test whether admission officers are more likely to admit a low-SES student if they have high-quality contextual information about median test scores in the student's high school and zip code. In addition, we examine whether admission officers are more likely to admit an applicant who demonstrates grit and whether the potential impact of grit varies by students' socioeconomic background by supplying an essay that demonstrates perseverance over obstacles. The findings of this study can contribute to researchers' understanding of the role of information in the application process and can inform institutional efforts to improve socioeconomic diversity.

## **Background**

### **Selective College Admissions Policy in the United States**

Colleges and universities across the United States receive millions of student applications each year. In an era when online applications have become nearly ubiquitous (Clinedinst & Koranteng, 2018), first-time freshman students are applying to a greater number of colleges than ever before (Eagan et al., 2014). Among the nation's most selective institutions, the relative competition for admission among students has steadily increased over time (Bastedo & Jaquette, 2011). As growing number of students strive to gain entry to selective colleges and universities in the United States, questions surrounding who gets admitted and why are important to both the academic community and society at large.

The goal of a college admission office, broadly speaking, is to admit candidates who are likely to be academically successful at the institution, while attending to the enrollment goals of diversity, academic profile, and revenue generation (Cheslock & Kroc, 2012). Institutions report a wide array of practices for reading and evaluating student applications, from highly formulaic processes that rely on traditional measures of high school grades and standardized test scores to holistic approaches that consider an applicant's academic and personal accomplishments within the context of their school and family environment (Bastedo, Bowman, Glasener, & Kelly, 2018). In the case of selective college admissions, where student demand for seats exceeds supply, institutions have substantial latitude in the kinds of students they admit. Selective colleges aiming to improve in the rankings, for instance, may place greater emphasis on standardized test scores in admission decision making, whereas resource-constrained colleges may implement admission procedures that foster the enrollment of revenue-generating students (Jaquette, Curs, & Posselt, 2016).

### **The Role of College Entrance Exams in Selective Admissions**

Although colleges and universities have adopted a variety of admission procedures to fit their institutional needs and priorities, consistent across most selective, 4-year institution is a strong emphasis on standardized test scores in the evaluation of applicants' academic qualifications (Clinedinst & Koranteng, 2018). Since first introduced in the early 20th century, standardized testing has seen extraordinary growth. Fewer than 1,000 students sat for the first College Board exam in 1901, for example, but by the late 1950s, more than half a million students took the SAT annually. Today, more than 1.8 million students take the SAT and/or ACT annually. In addition, 21 states now require students to take the SAT or the ACT, and three others give students a choice of taking the SAT or the ACT as part of their statewide graduation requirements.

The growing importance of college entrance exams is also reflected in admission policy; approximately 90% of the 4-year colleges and universities now require either the SAT or ACT tests for admission. Since 1993, the percentage of college admission officers reporting college entrance exams to be “considerably important” has increased by nearly 15 percentage points (Clinedinst & Koranteng, 2018). Although the number of schools who say test scores are “considerably important” has declined slightly over the past decade, the most recent National Association for College Admission Counseling (NACAC) Admission Trends Survey indicates that test scores are still the fourth-most important factor considered in admissions, behind grades in college prep courses, strength of curriculum, and overall grades (Clinedinst & Koranteng, 2018).

Although grades are often a stronger predictor of student success, standardized test scores provide a common and efficient measure to compare students from different high school and

geographic contexts. Because grading standards, instructional resources, and levels of rigor vary widely across high schools, using high school grade point average (GPA) alone as an admission indicator may mask differences in educational contexts and lead to errors in collegiate performance predictions for some student populations (Zwick & Himelfarb, 2011). College entrance exam scores, in contrast, arguably measure students' achievement on a test that is standardized across contexts and combined with high school GPA, serve as useful predictors of college performance.

The emphasis on entrance exam scores in admission decisions is also due in part to the methodologies of college ranking systems. Rankings such as U.S. News & World Report's America's Best Colleges consider the average test scores of the incoming class as a key measure of quality, representing 65% of the selectivity score. Researchers have found that improvement in U.S. News & World Report (USNWR) rankings is associated with greater number of applications, higher yield rates, and an increase in average freshman SAT/ACT scores (Bowman & Bastedo, 2009). In addition, the perception of selectivity is particularly valued among higher education administrators, and enrollment managers therefore face substantial pressures for the continued use of college entrance exams in admissions to ensure the prestige and reputation of their institutions.

### **Admission Disadvantages for Low-Income Students**

The SAT and ACT have been widely criticized as an impediment to low-SES students' access to selective colleges and universities (Atkinson & Geiser, 2009). Some scholars have called the predictive validity of standardized test scores into question, citing methodological concerns (Rothstein, 2004) as well as the misalignment of skills necessary for success in college and those measured by standardized test scores (Aguinis, Culpepper, & Pierce, 2016). However,



there is some evidence that the association between students' socioeconomic background and standardized test scores is not unique. Both SAT scores and high school GPA are highly correlated with SES when examined at the neighborhood (rather than collegiate) level (Zwick & Green, 2007). Moreover, research by Sackett, Kuncel, Arneson, Cooper, and Waters (2009) indicates that the association between SAT scores and first-year college GPA remains even after controlling for SES, suggesting that test scores are not mere proxies for family wealth. As a result, many selective colleges have moved toward test-optional or test-flexible policies (Syverson, Franks, & Hiss, 2018), although it is contested whether those policies improve equity for low-income students (Belasco, Rosinger, & Hearn, 2015).

The persistent socioeconomic differences in standardized test scores reflect important disparities in home and school resources (Grotsky, Warren, & Felts, 2008). Students from higher-SES households are typically segregated in higher-income neighborhoods, whereas those from low-income households are less frequently segregated in economically similar neighborhoods (Reardon, Fox, & Townsend, 2015). Neighborhood differences in household income are associated with differences in access to quality K-12 education (including course offerings), as well as differences in average standardized test scores (Zwick & Green, 2007). Higher-SES students can afford expensive test preparation services, such as private tutors and classes, that are often unavailable to low-SES students and are significantly more likely than their low-SES peers to take entrance exams multiple times in an attempt to increase their scores (Buchmann, Condron, & Roscigno, 2010; Vigdor & Clotfelter, 2003).

### **College Admissions and Correspondence Bias**

When admission officers consider college entrance exam scores in the admission process, they may implicitly assume that differences they observe in academic performance between

applicants are the result of differences in academic potential without considering differences in students' home and school environments or their access to quality secondary education. This assumption can be a form of correspondence bias, or the tendency of people to attribute causes to individual dispositions rather than to the situation in which actions occur (Gilbert & Malone, 1995; Ross & Nisbett, 2011). When reviewing applications, an admission officer may attribute a low-SES student's lower standardized test scores as a signal of ability or potential rather than the student's high school environment, which may lack resources such as access to advanced curricula or test preparation materials. Providing more robust contextual information to situate students' scores can reduce correspondence bias by allowing the admission officer to assess the student's scores relative to their immediate peers instead of the entire applicant pool.

In a holistic review process, standardized test scores are treated as one component of a student's application, and admission officers are instructed to consider a student's background when evaluating information and making decisions (Rigol, 2003). Nonetheless, correspondence bias may lead admission officers to attribute academic performance to a student's intelligence, ability, or motivation rather than the opportunities afforded by their family, high school, or neighborhood. This tendency can be shaped by the type of information provided in an application file and the form in which it is provided. For example, when admission officers receive more detailed information about a student's high school (e.g., number of Advanced Placement [AP] courses offered, percentage of students eligible for free/reduced lunch), they are significantly more likely to admit a low-SES applicant (Bastedo & Bowman, 2017).

Similarly, contextual information about a student's standardized test score relative to peers could shift admission officers' decisions. For a low-SES student whose standardized test scores are above the high school median, contextual information could result in an increased

likelihood of admission, as the information shows the student as having performed well relative to peers. For a high-SES student with similar test scores in a school with higher overall SAT scores, we would not expect to see any boost in the probability of admission because the student is performing closer to the average of the school. We hypothesize the following:

**Hypothesis 1a (H1a):** Admission officers will be more likely to recommend admission for high-performing low-SES applicants when provided with more detailed contextual information about test scores.

**Hypothesis 1b (H1b):** In contrast, admission officers will not be more likely to recommend admission for high-performing higher-SES applicants when provided with more detailed contextual information about test scores.

### **Potential Role of Non-Cognitive Factors in Admissions**

Although academic credentials are given the greatest weight in the admission process, some colleges consider students' personal characteristics and noncognitive factors as part of a holistic review process (Hossler et al., 2019; Rigol, 2003). Because selection based on test scores alone potentially decreases racial/ethnic and socioeconomic diversity and underestimates the likelihood of success for many capable students, institutions have a growing interest in considering non-cognitive predictors of college success (Akos & Kretchmar, 2017; Sternberg, 2010). The underlying assumption, as with reliance on standardized test scores, is that a positive association exists between the presence of positive non-cognitive factors and a student's likelihood of academic success. This assumption is reinforced by findings that perseverance of effort is a useful predictor of outcomes such as college GPA and intent to persist (Bowman, Hill, Denson, & Bronkema, 2015).

To some extent, students' non-cognitive characteristics have always been part of admission review, as most selective universities require essays, letters of recommendations, and a description of extracurricular activities, all of which give insights into an applicant's

personality and character. However, over the past two decades, new conceptualizations of non-cognitive traits have entered education and admission discussions, particularly related to the concepts of “grit” and “growth mind-set” (Duckworth, Peterson, Matthews, & Kelly, 2007; Dweck, 2006; Hossler et al., 2019). Grit relates to an individual’s ability to persevere in the pursuit of long-term goals and maintain consistent interest over time (Duckworth et al., 2007), whereas growth mindset speaks to the belief that one’s own talents and abilities can be developed (Dweck, 2006). These concepts have been applied and debated widely, with charter school networks like Knowledge Is Power Program (KIPP) and national remedial education initiatives such as Carnegie Math Pathways incorporating the ideas into their curricular design (Tough, 2011). The increased attention in education to these concepts, particularly grit, builds on admission offices’ use of holistic review by introducing another aspect that can be used to assess an applicant’s file. The challenge is in identifying and measuring whether a student has demonstrated grit or other potentially relevant non-cognitive factors.

Large-scale efforts have been directed at developing and evaluating new non-cognitive predictors of students’ academic performance, persistence, and degree attainment. For example, Sternberg, Bonney, Gabora, and Merrifield (2012) introduced a multi-dimensional construct of intelligence for college admissions inclusive of creativity, wisdom, and analytical and practical intelligence, instead of a one-dimensional measure of general ability as it is typically conceived. This measure demonstrated incremental validity in the prediction of GPA, above and beyond standardized test scores and high school GPA. Other institutions report the use of the Non-Cognitive Questionnaire (NCQ; Tracey & Sedlacek, 1986) or the Rainbow Project (Sternberg & Collaborators, 2006) as alternative admission devices. Knowledge of admission officers’ ongoing practices related to non-cognitive assessment is largely anecdotal. For many years,

campuses in the University of California system have used “augmented review” to examine how students persist in the face of substantial life obstacles, and these students are primarily identified through their personal statements (Hout, 2005; Mare, 2012). Based on the increase in popularity of concepts like “grit” as well as more general interest in non-cognitive factors in the admission process, we hypothesize that, regardless of SES status:

**Hypothesis 2 (H2):** Admission officers will be more likely to recommend admission for applicants that exhibit grit or perseverance of effort in their essays.

## Method

### Participants

We drew upon a pool of college admission officers who work at a selective college or university, defined as top three tiers of Barron’s (2013) competitiveness ratings. Participants were recruited from attendees of the 2014 annual meeting of NACAC; we limited invitations to those whose job title implied that they would regularly review applications (e.g., admission counselor and director of admissions). Recruitment emails were sent to 960 conference attendees who met the inclusion criteria (34% overall response rate). Admission officers who agreed to complete the simulated admission file review and survey received US\$50 for their participation. Among the admission officers who participated, 57% were female, 75% were White/Caucasian, 8% were Black/African American, 6% were Latino/Hispanic/Chicano, 3% were Asian American/Pacific Islander, 1% were American Indian/Native American, and 7% were from multiple racial/ethnic groups. The extent to which these participants are representative of college admission officers nationally is unclear, as no census data on selective college admission officers exist.

A total of 350 admission officers clicked on the survey invitation, but only 321 provided recommendations for all four simulated applications (these participants generally dropped out

before reviewing a single admission file). Attrition was fairly evenly distributed between treatment and control conditions (13 and 16 participants, respectively). Of these 321 participants, an additional 15 did not respond to all the demographic and admission office questions; therefore, just less than 5% of the sample had at least one missing value, and 0.5% of all values for the predictors were missing. We used multiple imputation to analyze the incomplete data; this technique minimizes bias and maximizes—but does not overestimate—precision when compared with listwise deletion and other approaches (Little & Rubin, 2002). Specifically, we created 50 datasets via multiple imputation by chained equations to account appropriately for the binary or ordinal nature of most predictors (as described below). All independent variables described below were used to impute one another. Test score contextual information and institutional selectivity had complete data, whereas the majority of the remaining variables had only one participant whose response was missing. The analyses were conducted for each of the 50 datasets; the parameter estimates were averaged across datasets, and the standard errors included variation that occurred both within and across datasets. Preliminary analyses using listwise deletion yielded the same general pattern of results; if anything, the multiple imputation analyses provided in this article yielded slightly more conservative estimates than those from listwise deletion.

### **Survey Development and Procedure**

Admission officers were first asked to review four simulated admission files. Following the experimental component, participants answered a brief survey that asked about their admission office practices, as well as their own background and professional experience. The survey was administered online using Qualtrics survey software. Each participant was provided with application files that were tailored to the Barron's selectivity tier where they work so that

applicants would have a reasonable chance of acceptance or rejection. Specifically, we used data from Education Longitudinal Study (ELS:2002) to inform the creation of simulated admission files. ELS is a nationally representative dataset that surveyed more than 15,000 10th graders in 2002, with follow-up surveys in 2004, 2006, and 2012. This dataset contains information on student demographics, high school academic performance, and postsecondary applications and enrollments. Using this information, we computed the deciles for high school GPA (weighted and unweighted), SAT test scores, and AP coursework for students who applied to institutions in each selectivity tier and used this information to craft simulated applications that would be competitive in each tier. Because more selective schools reject a larger percentage of applicants, we made the percentiles for applicants' credentials higher for more selective tiers, whereas the credentials for applicants in the less selective tiers reflected a lower ELS percentile as these institutions have higher acceptance rates.

We also assigned the simulated files for low-SES applicants more modest SAT scores in each tier relative to the high-SES applicants, as low-SES students tend to achieve lower standardized test scores, on average, compared with their more affluent counterparts (Zwick, 2004). The unweighted GPAs of low-SES applicants were higher than those of their higher-SES peers to reflect the fact that these students were performing well in the courses available to them. However, weighted GPAs for the high-SES applicants were higher because of the greater availability of honors and AP courses in their high schools. Finally, to test whether contextual information regarding low-SES students' test scores influenced admission ratings and recommendations, the simulated files contained raw SAT scores that significantly exceeded their median high school and zip code test scores.

Within the online survey, each section of the application was presented on a separate page, and participants could go back to earlier pages if they desired. Each application contained information about the student's high school, including the high school name (fictitious), state (all from the northeast region of the United States), institutional control (e.g., public/private), and number of students in the high school. The files included students' academic qualifications (i.e., unweighted and weighted high school GPA, number of honors/AP/International Baccalaureate [IB] courses taken, AP examinations and scores, and scores for each sub-section of the SAT), extracurricular activities, and personal statement. The admission file also contained the applicants' parental education, so participants knew at least one dimension of applicants' SES. They were also provided with the overall high school graduation rate, which is especially important because graduation rates are strongly associated with the average SES of students at the high school (Freeman & Simonsen, 2015). All parents of the higher-SES applicants had at least a master's degree, and these students attended high schools with graduation rates above 95%. In contrast, the low-SES applicants had parents with education levels less than a bachelor's degree (e.g., less than high school diploma, high school diploma, or associate's degree) and attended high schools with graduation rates around 65%. Therefore, all participants received some important contextual information about the applicants and their high schools.

Because admission recommendations can vary notably depending upon the race/ethnicity and gender of applicants, and the college or major to which they apply, these attributes were held constant across applications; all four students were White females who listed social sciences (i.e., psychology, sociology, anthropology, or political science) as their intended major. To illustrate low- and high-SES applicants' relative qualifications, an overview of their academic indicators



for the most competitive tier, along with high school and parental education information, is provided in Table 1.

TABLE 1

**Overview of Applicants' Academic and School Indicators for the Most Competitive Tier (in SAT Context Condition)**

[Insert Table 1 here]

NOTE: AP = Advanced Placement; SES = socioeconomic status

NOTE: Horizontal bars distinguish between academic and high school information as well as parental education, which appears in all conditions. The high school median and zip code median SAT scores only appear in the SAT context condition. Weighted HSGPA was computed by adding ½ point for an honor class and a full point for an AP class; a description of the weighting approach was provided to participants within the simulated application files.

We conducted two experimental manipulations—one between-subject and one within-subject—to examine whether providing contextual information about test scores and overcoming obstacles influenced admission recommendations. Figure 1 provides a visual representation of the experimental manipulations. The first experimental manipulation was across subjects: admission officers were randomly assigned to treatment and control, where treatment participants' applications all included median SAT test scores for the applicants' high school and zip code. For experimental purposes, the high school and zip code averages were essentially the same for each applicant; in other words, no applicant attended a high school whose average differed notably from that of their zip code. In the contextual (i.e., treatment) condition, participants ( $n = 162$ ) received the applicant's own test scores (critical reading, mathematics, and writing) as well as the corresponding median SAT scores for the high school and zip code in all four of the applications they reviewed. Participants in the limited-contextual condition ( $n = 159$ ) only observed applicants' own SAT test sub-scores.

## FIGURE 1

**Experimental Randomization**

[Insert Figure 1 here]

NOTE: SES = socioeconomic status.

The second experimental manipulation was whether the admission officer reviewed a personal essay that portrayed an important form of “grit.” Each participant reviewed two applications (one high-SES and one low-SES) with “gritty” essays and two applications (one high-SES and one low-SES) with a standard personal statement. Thus, the student essay manipulation served as a within-subject experimental manipulation. The sequence in which the applications appeared was counterbalanced across participants to remove the confounding influence of potential order effects.

Duckworth et al.’s (2007) conceptualization of grit consists of two related dimensions: perseverance of effort and consistency of interest. Given that perseverance of effort is the much stronger and more consistent predictor of college student success (e.g., Bowman et al., 2015), these essays illustrated this component through a student discussing how she successfully overcame a challenge. Specifically, one applicant described what it was like growing up in a home without a father and with a mother who had to work long hours to support her family, whereas another shared her experience as the daughter of an alcoholic. In both essays, the applicants conveyed how they had overcome adversity, highlighting the skills they had developed because of their situations. The non-grit essays, on the contrary, conveyed students’ personal experiences and passions, but did not specifically emphasize obstacles or personal challenges that demanded fortitude or perseverance. All essays were closely based on real college essays that students had submitted for admission to selective colleges and universities.

This research was supported in part by grant R305B150012# from the Institute of Education Science at the U.S. Department of Education.

These essays were made publicly available by selective colleges themselves and were retrieved from the Internet. Minor tweaks (e.g., gender and geographic region) were made for consistency with other aspects of the simulated applications.

As a supplemental analysis, we sought to obtain external validation to test whether two of these essays exhibited notably greater perseverance of effort than the others. Therefore, we recruited seven admission officers at selective colleges and universities; these participants worked at a mixture of large research universities and small liberal arts colleges. We provided these admission officers with the four essays and asked them to rate each one in terms of perseverance of effort using Duckworth et al.'s (2007) original six-item scale for this construct ( $\alpha = .81$ ). As expected, participants rated the two essays that we believed were "gritty" as displaying notably higher levels of perseverance of effort than the other two essays (Cohen's  $d = .94$ ). Although this small sample was intended to serve only as a pilot study, a paired t-test showed that this difference across the two types of essays was statistically significant,  $t(6) = 2.477, p < .05$ .

After reading each subsection of a given application, participants were asked to provide ratings of the quality of academic record, extracurricular activities, and personal statement. Once the participant had read the entire application, they provided the admission recommendation (deny, wait list, or accept) that they would make if that applicant had applied to the institution at which they work. The realism of this rating process was facilitated by participants using the same criteria than they normally use at their institution and by not being forced to pick one applicant over the other (they did not have a limit on the specific number of acceptance, wait list, or rejection recommendations). We collapsed the recommendations into a dichotomous measure to capture acceptance recommendations. Several of the ordinal logit analyses did not meet the

parallel line assumption, so it was not appropriate to use this technique (e.g., Long, 1997); moreover, the substantive conclusions are similar using binary and ordinal outcomes. From a practical standpoint, many students who are placed on wait lists at selective institutions are never ultimately accepted (Clinedinst & Koranteng, 2018), so using this acceptance outcome may be more meaningful.

### **Measures**

The primary dependent variable was a binary measure indicating participants' admission recommendation for each applicant (0 = deny or wait list, 1 = accept). We also examined participants' ratings of each applicant's academic record, extracurricular activities, and personal statement (1 = very poor to 6 = excellent). The main independent variables of interest indicated the between-subject experimental manipulation (0 = no SAT context, 1 = SAT context) and the within-subject manipulation (0 = low-grit essay, 1 = high-grit essay). In addition, we included dummy-coded variables to capture selectivity tier, with the most competitive institutions as the referent group.

We included several additional covariates in the model, including participants' sex (0 = male, 1 = female), race/ethnicity (given the small sample sizes for some groups, a single dichotomous indicator was used in which 0 = White/Caucasian, 1 = participant of color), parental education (1 = elementary school to 9 = graduate degree), experience working in admissions (1 = less than 1 year to 7 = 21 years or more), and whether they were working at the same institution from which they received their bachelor's degree (0 = no, 1 = yes). We also included attributes of the admission offices in which participants worked to determine whether specific policies or practices were associated with acceptance decisions. Specifically, binary variables (0 = no, 1 = yes) were used to indicate each of the following admission office

attributes: whether grades are the most important criterion for determining academic merit; whether test scores are the most important criterion for determining academic merit; whether the participant considers standardized test scores differently if the student is from a low-income family or attended an underresourced school; whether high school information is consistent across all applicants; whether the participant is required to write a paragraph explaining their admission recommendation; and two separate measures indicating whether committees consider (a) students' demonstrated interest and (b) potential fit with the institution as part of the review process. Participants rated the average quality of high school information that they receive on a 6-point scale (1 = poor, 6 = excellent), reflecting an approximate measure of how often they receive high school profiles and the quality of the information included in the profiles they receive. Finally, participants rated the extent to which they consider students' non-cognitive attributes in their assessment of an admission file (1 = not at all, 4 = a great deal).

Table 2 provides descriptive statistics for the SAT context and no context conditions. Although participants were randomly assigned, admission officers in the SAT context were significantly more likely to report that the high school information they receive in their admission office is similar across applications and that they consider applicants' demonstrated interest when making recommendations. As described below, the analyses controlled for these two variables to correct for these group differences, and additional covariates were also included to reduce error variance (Raudenbush, Martinez, & Spybrook, 2007).

TABLE 2

**Summary Statistics by SAT Context and Control (Mean Values and Standard Deviations)**

[Insert Table 2 here]

NOTE: Chi-square and one-way ANOVA tests were conducted on all covariates to examine differences in mean values across treatment and control groups. ANOVA = analysis of variance.

\* $p < .05$

This research was supported in part by grant R305B150012# from the Institute of Education Science at the U.S. Department of Education.

## Analysis

The initial analyses examining the effect of the between-subject manipulation of test score contextual information predicting acceptance recommendations were conducted separately for each simulated applicant. Given the dichotomous nature of the dependent variable, logistic regression analyses were used to predict acceptance recommendations (Long, 1997). The statistical equation for predicting the binary acceptance recommendation is the following:

$$\ln \frac{p(\gamma_{ij})}{1 - p(\gamma_{ij})} = \beta_0 + \beta_1 \text{SAT CONTEXT}_j + \beta_2 \text{SELECTIVITY}_j + \beta_3 \text{participant characteristics}_j + \beta_4 \text{office characteristics}_j + e_{ij} \quad (1)$$

The variance inflation factors for all variables were below 2.3, so multicollinearity did not appear to be a concern.

The logistic regression analyses can determine the effect of the between-subject random assignment of SAT contextual information on acceptance recommendations. In fact, these separate regression analyses are necessary to determine the effect for each individual applicant, which thereby sheds light into Hypotheses 1a and 1b (whether the experimental condition affects acceptance recommendations for low-SES and high-SES applicants, respectively). This type of “single-level” analysis is appropriate here because only one applicant is being examined at a time, so the assumption of independence of observations is not violated.

However, such analyses are not suitable for the within-participant essay manipulation, because the comparison of applicants with gritty essays versus those with less gritty essays inherently requires using multiple ratings from the same participant simultaneously. Therefore, multilevel logistic regression analyses were also conducted with applicant recommendations (level 1) nested within participants (level 2; Raudenbush & Bryk, 2002). Considerable variation in acceptance outcomes occurred across participants, as indicated by the intraclass correlation (ICC) coefficient (.76). The ICC is well above the suggested value of .05 that typically

necessitates multilevel modeling (Heck & Thomas, 2009). In the multilevel analyses, all predictors used in the regression analyses were included at level 2, and binary variables indicating whether the personal essay portrayed grit and whether the applicant was from a low-SES background were included at level 1. Additional analyses incorporated interactions between SAT context and these two within-participant predictors. Finally, ordinary least squares multiple regression analyses were conducted to predict ratings of the intermediate outcomes for each applicant: academic qualifications, extracurricular activities, and personal statements. The threshold for establishing statistical significance of findings was  $p < .10$ , as the modest sample size limits the statistical power of the analyses, and this intervention is low-cost and low-risk (so the concerns about Type I error are somewhat reduced).

### **Limitations**

The primary limitation of this study is that we were unable to conduct a field experiment in real-world conditions, instead relying on a simulation of admission decision making. As a result, it is possible that admission officers may not make the same scoring and admission recommendations they would make if their decisions had real consequences. Admission officers may be more likely to admit low-SES applicants in the low-stakes environment of the simulation. That said, the hypotheses of the study do not address the relative likelihood of admissions for low- versus high-SES applicants overall; instead, they suggest that the effect of SAT context will differ depending upon the SES of the applicant. Thus, to provide an alternative explanation for the hypotheses, this leniency would need to operate only for low-SES students when SAT context is provided (this explanation cannot account for all of the SAT context effects that we actually observed, as described below).

Furthermore, at many colleges or universities, admission officer recommendations may not result in formal acceptance. Reader recommendations are often taken to an admission committee for further review, and in some cases, final admission decisions are determined by deans or directors of enrollment management. Low-SES applicants who receive high admission scores and recommendations in this experiment may ultimately not be admitted once final decisions are made. The results presented here on admission scoring should therefore be interpreted as recommendations rather than final admission decisions. Nonetheless, it is important to remember that these factors are most likely to affect all decisions in the simulation regardless of the condition to which they were assigned.

Because this experiment was a simulation, it is also possible that participants might try to determine the aim of the study and therefore give answers that they thought we were looking for. To mitigate this risk, participants were told that the study only sought to better understand admission decision making, and the study materials did not indicate that the analytic focus was on the socioeconomic backgrounds of students or test scores in context. Participants were instructed to read each application as a “real file” for their school, not in some hypothetical context, which further bolsters the generalizability and applicability to real-world decision making. Moreover, the applicants’ sex (female), race (White), geographical region (northeast), and general field of study (social sciences) were held constant so that these extraneous aspects of the file did not confound the results. Previous research has suggested that admission officers may give some preference to applicants with SES backgrounds that are similar to their own (Bowman & Bastedo, 2018), but participants’ identities did not moderate the impact of contextual information on admission recommendations (Bastedo & Bowman, 2017). In addition, we must also consider that even if admission officers are looking for non-cognitive skills or abilities, they



may not find it in an essay that they cannot triangulate with other crucial information, such as letters of recommendation or certain extracurricular activities.

## Results

### Hypothesis 1: Test Scores in Context

We begin by testing our first set of hypotheses, which predicted that providing contextual information for low-SES students' SAT scores will result in a higher probability of receiving acceptance recommendations from admission officers. Table 3 reports the coefficients for the effect of providing contextual information about students' SAT scores, both as an odds ratio and an average marginal effect, on admission recommendations. Consistent with H1a, admission officers were significantly more likely to recommend admission for the low-SES applicant with a high-grit essay when provided with more detailed contextual information about test scores. Results indicate that the average marginal effect for admitting this applicant in the contextual condition was around 9 percentage points ( $p = .07$ ), controlling for institutional selectivity, participant characteristics, and admission office attributes. Providing test score context also contributed to more positive assessments of this applicant's academic qualifications (see Table 4). The effect of SAT context on acceptance recommendations was similar in magnitude for both low-SES applicants; however, adding contextual information about test scores did not result in a significant difference in the likelihood of acceptance for the low-SES applicant with a low-grit essay ( $p = .12$ ). Chi-square and logistic regression analyses without control variables showed a significant, positive effect of SAT context on acceptance recommendations for the low-SES, low-grit applicant ( $p = .09$ ). No significant results were apparent for the academic, extracurricular, or essay ratings.

TABLE 3

**Logistic Regression Results and AMEs for Admission Recommendations ( $n = 321$ )**

[Insert Table 3 here]

NOTE: Acceptance recommendations were coded as 0 = deny or wait list, 1 = accept. Coefficients presented are odds ratios; standard errors are in parentheses. AME = average marginal effects; SES = socioeconomic status. † $p < .10$ . \* $p < .05$ .

TABLE 4

**Results for Multiple Regression Analyses Predicting Ratings of Academic Record, Extracurricular Activities, and Personal Statements in SAT Context Condition ( $n = 321$ )**

[Insert Table 4 here]

NOTE: Academic record, extracurricular activities, and personal statements were all coded on a 6-point scale (1 = *very poor* to 6 = *excellent*). Standard errors are in parentheses. SES = socioeconomic status. † $p < .10$ . \* $p < .05$ .

Contrary to H1b, providing context about SAT scores also significantly improved the acceptance recommendations and ratings of academic qualifications for the high-SES applicant with a less gritty personal statement. Results indicate that the average marginal effect for admitting this student in the contextual condition was around 11 percentage points ( $p = .06$ ). Participants in the SAT context condition also rated this applicant's academic qualifications 0.14 points higher than those in the control condition, net of controls ( $p = .06$ ). For the high-SES, high-grit applicant, test score context did not affect acceptance recommendations, but it did predict higher ratings of extracurricular activities and the personal statement (see Table 4). Among the control variables, the most consistently significant predictors were institutional selectivity and use of test-optional admission policies. Participants at Tier 3 institutions were more likely to recommend accepting all applicants than those at Tier 1 institutions, and test-optional policies were positively related to acceptance recommendations for all but the low-SES,

high-grit applicant. Overall, providing context about SAT scores increased the probability of an admission recommendation and the ratings of individual file components for low-SES applicants, supporting the argument that providing scores in context produces more favorable admission probabilities. However, because these effects were similar in magnitude across applicants, applicants from low socioeconomic backgrounds did not differentially benefit from providing more robust contextual information about test scores in this experiment.

## **H2: Non-Cognitive Factors and College Admissions**

We examined through multiple types of statistical tests whether admission officers were more likely to provide favorable admission reviews to students who demonstrate the ability to overcome adversity. The admission files of same-SES peers were designed to have very similar qualifications between the high-grit and low-grit files, so any differences in outcomes could be attributed to the grittiness of the personal statement. First, repeated-measures analyses of variance showed no significant differences in acceptance recommendations or ratings of any aspect of the admission file across the four applicants ( $ps > .29$ ). Next, multilevel logistic regression analyses explored the overall effect of providing SAT context on acceptance decisions and how these effects might vary across applicants with differing levels of grit as exhibited in the personal statement. As shown in Table 5, admission officers were more likely to recommend acceptance when receiving additional context about applicants' test scores. Specifically, Model 1 provides results for the effect of context across ratings of all applicants; the odds of recommending acceptance are 2.44 times when SAT context is present than when it is not, net of controls. Additional analyses that contained interaction terms (Model 2) showed that these context effects did not differ significantly between low- and high-SES applicants. Multilevel logistic regression analyses also showed that applicants who portrayed grit in their personal

statements did not differ in acceptance recommendations than their lower-grit peers; there was also no significant interaction between essay grit and test score context. In short, the extent to which the personal statement expressed the applicant's "grittiness" was not significantly related to the probability of acceptance. In terms of demographic control variables, participants of color were less likely to recommend acceptance of these applicants on average, but no other participant attributes were significant predictors.

TABLE 5

**Results for Multilevel Logistic Regression Analyses Predicting Acceptance Recommendations (Level 1) Nested within Participants (Level 2) ( $n = 321$ )**

[Insert Table 5 here]

NOTE: Acceptance recommendations were coded as 0 = deny or wait list, 1 = accept. Coefficients presented as odds ratios; standard errors are in parentheses. Reference groups were tier 1 for selectivity and curriculum rigor as the most important merit criterion. SES = socioeconomic status.

† $p < .10$ . \* $p < .05$ . \*\* $p < .001$ .

### **Discussion and Implications**

Our results suggest that adding contextual information about test scores to the admission file influences admission decisions. Consistent with the philosophy of holistic review, admission officers showed a willingness to reward lower-SES applicants for performing above average for their schools and neighborhoods on the SAT. Admission officers can use this information to better evaluate students' test scores by contextualizing their raw scores with average school and neighborhood test score performance. For selective colleges seeking to improve the socioeconomic diversity of their student bodies, adding contextual information on SAT performance for low-SES students in particular may be an effective way to improve admission ratings and acceptance of such students. There is also strong evidence that participants read and

evaluated the applications in context because they rated some individual components of the file more favorably in the treatment condition.

However, our findings also indicate that providing additional context about students' SAT scores leads to more positive evaluations, regardless of socioeconomic background, suggesting that such information may not serve to "level the playing field" between higher- and lower-SES students in the admissions process. Admission officers may give preference to applicants from high schools with high average SAT scores, regardless of how they compare with their similarly situated peers. This may even be rational, given research showing that the average score at a high school may be more predictive of college GPA than a student's own raw SAT score (Rothstein, 2004). However, given unequal educational opportunities and the fact that SAT scores are highly correlated with a student's SES, this type of selection behavior overwhelmingly benefits economically advantaged students who attend highly resourced high schools where average SAT scores tend to be higher, thus contributing to a concentration of wealthy students at selective colleges and universities. To mitigate this risk and increase access for high-performing, low-income students, admission officers must "strive to understand the conditions under which the applicant has performed and to make judgments based on the context of those conditions" (Lucido, 2014, p. 157). Therefore, it may be particularly important to educate and train individuals entering the field of college admissions about the validity of standardized test scores, and the persistent relationship between students' racial/ethnic and socioeconomic background and SAT scores.

In addition, this study has important implications for understanding how non-cognitive factors such as grit influence admission decision making. Although many low-SES students are advised to write personal essays that convey their backgrounds and diversity of experiences, our

findings suggest that such essay topics may not lead to more positive evaluations in the selective admission process. Student's personal characteristics and non-cognitive factors are often conveyed in multiple components of the application. It is possible that admission officers rely more heavily on other aspects of the application, such as letters of recommendations, community service, and extracurricular activities (Thomas, Kuncel, & Credé, 2007), to evaluate applicants' non-academic qualities. The simulated admission files included a record of students' extracurricular activities during high school, but they did not include teacher or counselor recommendation letters.

It is also possible that participants may not have found the essays provided in the experiment to be effective in conveying students' perseverance and other non-cognitive strengths, despite the fact that the "grit" essays delved into overcoming substantial obstacles. Non-cognitive traits like grit may also be better captured in a student's application through formal questionnaires rather than informally through an applicant's personal statement. Several scales have been developed to measure a student's "grittiness," which have been found to predict markers of success in college (Bowman et al., 2015; Duckworth et al., 2007). If admission offices are interested in better integrating non-cognitive evaluations into their evaluation process, then they may be better served by introducing more formal assessments of such personal characteristics as grit, determination, creativity, or work ethic. However, self-report scales, like those measuring grit, have not yet been validated for use under high-stakes conditions such as college admissions because of susceptibility to social desirability response bias (Akos & Kretchmar, 2017), and grit may not serve as an independent psychological construct from perseverance of effort or conscientiousness (Credé, Tynan, & Harms, 2017). Rather than having applicants complete a separate survey to measure noncognitive traits, college admission offices

may consider relying on letters of recommendation from teachers and counselors to assess these student characteristics.

This study's findings have substantial implications for decision making in selective college admission offices. Today, admission officers often have low-quality information about students' high school and family contexts when evaluating applications (Bastedo & Bowman, 2017). They are also often underinformed about differences in SAT score performance by school, neighborhood, or even region of the country. Even when admission officers know they should account for contextual information, normal human biases ensure that they will often fail to do so without sufficient information (Gilbert & Malone, 1995). In addition to providing the median test scores for students' neighborhood and school, it may be useful to explore alternative ways of displaying contextual information regarding students' test scores.

Future research is needed to extend and expand this inquiry. Access to applicant-level data is essential to disentangle the exact barriers for low-SES students and mechanisms underlying admission decision making, especially if the goal is to increase access for low-SES students. Without such data, researchers must instead rely on experiments that simulate admission decision making. Examining longitudinal data from multiple institutions would provide stronger information on patterns of college admission decision making. More research is needed to understand both what information should be provided to admission officers to better contextualize their reading and interpretation of applicants' files and the most effective format to deliver such information to admission officers. Future studies might explore how admission officers are trained and instructed to read files, which could provide more insights into the "black box" of admission decision making. In addition, certain admission office practices, such as the use of maximum/minimum test score and GPA cutoffs and subjective evaluation of students'

personal essays, may affect decision making to a greater degree than many admission leaders and enrollment managers are aware.

### **ORCID iDs**

Michael N. Bastedo <https://orcid.org/0000-0003-3358-2564>

Nicholas A. Bowman <https://orcid.org/0000-0001-8899-7383>

### **References**

- Aguinis, H., Culpepper, S. A., & Pierce, C. A. (2016). Differential prediction generalization in college admissions testing. *Journal of Educational Psychology, 108*, 1045-1059.
- Akos, P., & Kretchmar, J. (2017). Investigating grit as a non-cognitive predictor of college success. *The Review of Higher Education, 40*, 163-186.
- Alon, S., & Tienda, M. (2007). Diversity, opportunity, and the shifting meritocracy in higher education. *American Sociological Review, 72*, 487-511.
- Atkinson, R. C., & Geiser, S. (2009). Reflections on a century of college admissions tests. *Educational Researcher, 38*, 665-676.
- Barron's. (2013). *Profiles of American colleges 2013* (30th ed.). Hauppauge, NY: Barron's Educational Series.
- Bastedo, M. N., & Bowman, N. A. (2017). Improving admission of low-SES students at selective colleges: Results from an experimental simulation. *Educational Researcher, 46*, 67-77.
- Bastedo, M. N., Bowman, N. A., Glasener, K. M., & Kelly, J. L. (2018). What are we talking about when we talk about holistic review? Selective college admissions and its effects on low-SES students. *The Journal of Higher Education, 89*, 782-805.
- Bastedo, M. N., & Jaquette, O. (2011). Running in place: Low-income students and the dynamics of higher education stratification. *Educational Evaluation and Policy Analysis, 33*, 318-339.



- Belasco, A. S., Rosinger, K. O., & Hearn, J. C. (2015). The test-optional movement at America's selective liberal arts colleges: A boon for equity or something else? *Educational Evaluation and Policy Analysis, 37*, 206-223.
- Bowen, W. G., & Bok, D. C. (1998). *The shape of the river*. Princeton, NJ: Princeton University Press.
- Bowman, N. A., & Bastedo, M. N. (2009). Getting on the front page: Organizational reputation, status signals, and the impact of U.S. news and world report on student decisions. *Research in Higher Education, 50*, 415-436.
- Bowman, N. A., & Bastedo, M. N. (2018). What role may admissions office diversity and practices play in equitable decisions? *Research in Higher Education, 59*, 430-447.
- Bowman, N. A., Hill, P. L., Denson, N., & Bronkema, R. (2015). Keep on truckin' or stay the course? Exploring grit dimensions as differential predictors of educational achievement, satisfaction, and intentions. *Social Psychological and Personality Science, 6*, 639-645.
- Buchmann, C., Condron, D., & Roscigno, V. (2010). Shadow education, American style: Test preparation, the SAT and college enrollment. *Social Forces, 89*, 435-462.
- Cheslock, J., & Kroc, R. (2012). Managing college enrollments. In R. Howard, G. McLaughlin, & W. Knight (Eds.), *The handbook of institutional research* (pp. 221-236). San Francisco, CA: Jossey-Bass.
- Chetty, R., Friedman, J. N., Saez, E., Turner, N., & Yagan, D. (2017). *Mobility report cards: The role of colleges in intergenerational mobility* (W23618). Cambridge, MA: National Bureau of Economic Research.
- Clinedinst, M., & Koranteng, A.-M. (2018). *State of college admission 2017*. Washington, DC: National Association for College Admission Counseling.

- College Board. (2002). Best practices in admissions decisions: A report on the third College Board conference on admission models. New York, NY: Author.
- Credé, M., Tynan, M. C., & Harms, P. D. (2017). Much ado about grit: A meta-analytic synthesis of the grit literature. *Journal of Personality and Social Psychology*, 113, 492-511.
- Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, 92, 1087-1101.
- Dweck, C. S. (2006). *Mindset: The new psychology of success*. New York, NY: Penguin Random House.
- Eagan, K., Stolzenberg, E. B., Ramirez, J. J., Aragon, M. C., Suchard, M. R., & Hurtado, S. (2014). *The American freshman: National norms fall 2014*. Los Angeles: Higher Education Research Institute, University of California, Los Angeles.
- Frank, R. H., & Cook, P. J. (1995). *The winner-take-all society*. New York, NY: The Free Press.
- Freeman, J., & Simonsen, B. (2015). Examining the impact of policy and practice interventions on high school dropout and school completion rates: A systematic review of the literature. *Review of Educational Research*, 85, 205-248.
- Gilbert, D. T., & Malone, P. S. (1995). The correspondence bias. *Psychological Bulletin*, 117, 21-38.
- Grodsky, E., Warren, J. R., & Felts, E. (2008). Testing and social stratification in American education. *Annual Review of Sociology*, 34, 385-404.
- Heck, R. H., & Thomas, S. L. (2009). *An introduction to multilevel modeling techniques* (2nd ed.). New York, NY: Routledge.

- Hossler, D., Chung, E., Kwon, J., Lucido, J., Bowman, N., & Bastedo, M. (2019). A study of the use of nonacademic factors in holistic undergraduate admissions reviews. *The Journal of Higher Education*, 2019, 1-27.
- Hout, M. (2005). Berkeley's comprehensive review method for making freshman admissions decisions: An assessment. Berkeley: Academic Senate, University of California Berkeley.
- Hoxby, C. M., & Avery, C. (2012). The missing "one-offs": The hidden supply of high-achieving, low-income students (NBER Working Paper 18586). Cambridge, MA: National Bureau of Economic Research.
- Jaquette, O., Curs, B. R., & Posselt, J. R. (2016). Tuition rich, mission poor: Nonresident enrollment growth and the socioeconomic and racial composition of public research universities. *The Journal of Higher Education*, 87, 635-673.
- Lemann, N. (1999). *The big test: The secret history of the American meritocracy*. New York, NY: Farrar, Straus and Giroux.
- Little, R. J., & Rubin, D. B. (2002). *Statistical analysis with missing data*. New York: John Wiley & Sons.
- Long, J. S. (1997). *Regression models for categorical and limited dependent variables*. Thousand Oaks, CA: SAGE.
- Lucido, J. A. (2014). How admission decisions get made. In D. Hossler & B. Bontrager (Eds.), *Handbook of strategic enrollment management* (pp. 147-173). San Francisco, CA: Jossey-Bass.
- Mare, R. D. (2012, January). *Holistic review in freshman admissions at the University of California, Los Angeles*. Report prepared for the Committee on Undergraduate Admissions and Relations with Schools, University of California, Los Angeles.

- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks, CA: SAGE.
- Raudenbush, S. W., Martinez, A., & Spybrook, J. (2007). Strategies for improving precision in group-randomized experiments. *Educational Evaluation and Policy Analysis*, 29, 5-29.
- Reardon, S. F., & Bischoff, K. (2011). Income inequality and income segregation. *American Journal of Sociology*, 116, 1092-1153.
- Reardon, S. F., Fox, L., & Townsend, J. (2015). Neighborhood income composition by household race and income, 1990–2009. *The Annals of the American Academy of Political and Social Science*, 660, 78-97.
- Rigol, G. W. (2003). *Admissions decision-making models: How US institutions of higher education select undergraduate students*. New York, NY: College Board.
- Ross, L., & Nisbett, R. E. (2011). *The person and the situation*. New York, NY: McGraw-Hill.
- Rothstein, J. M. (2004). College performance predictions and the SAT. *Journal of Econometrics*, 121, 297-317.
- Sackett, P. R., Kuncel, N. R., Arneson, J. J., Cooper, S. R., & Waters, S. D. (2009). Does socioeconomic status explain the relationship between admissions tests and post-secondary academic performance? *Psychological Bulletin*, 135, 1-22.
- Shaw, E. J., Kobrin, J. L., Patterson, B. F., & Mattern, K. D. (2012). *The validity of the SAT for predicting cumulative grade point average by college major (College Board Research Report 2012-06)*. New York, NY: College Board.
- Sternberg, R. J. (2010). *College admissions for the 21st century*. Cambridge, MA: Harvard University Press.

- Sternberg, R. J., Bonney, C. R., Gabora, L., & Merrifield, M. (2012). WICS: A model for college and university admissions. *Educational Psychologist*, 47, 30-41.
- Sternberg, R. J., & Collaborators, T. R. P. (2006). The rainbow project: Enhancing the SAT through assessments of analytical, practical, and creative skills. *Intelligence*, 34, 321-350.
- Stevens, M. L. (2009). *Creating a class*. Cambridge: Harvard University Press.
- Syverson, S. T., Franks, V. W., & Hiss, W. C. (2018). *Defining access: How test optional works*. Arlington, VA: National Association for College Admission Counseling.
- Thomas, L. L., Kuncel, N. R., & Credé, M. (2007). Noncognitive variables in college admissions: The case of the non-cognitive questionnaire. *Educational and Psychological Measurement*, 67, 635-657.
- Tough, P. (2011, September 18). What if the secret to success is failure. *The New York Times*, p. 14.
- Tracey, T. J., & Sedlacek, W. E. (1986, April). Prediction of college graduation using noncognitive variables by race. Paper presented at the Annual Meeting of the American Educational Research Association (70th, San Francisco, CA).
- Vigdor, J. L., & Clotfelter, C. T. (2003). Retaking the SAT. *Journal of Human Resources*, 38, 1-33.
- Zwick, R. (2004). Is the SAT a “wealth test?” The link between educational achievement and socioeconomic status. In R. Zwick (Ed.), *Rethinking the SAT: The future of standardized testing in university admissions* (pp. 203-216). New York, NY: RoutledgeFalmer.
- Zwick, R., & Green, J. G. (2007). New perspectives on the correlation of SAT scores, high school grades, and socioeconomic factors. *Journal of Educational Measurement*, 44, 23-45.

Zwick, R., & Himelfarb, I. (2011). The effect of high school socioeconomic status on the predictive validity of SAT scores and high school grade-point average. *Journal of Educational Measurement*, 48, 101-121.

**Table 1:**  
**Overview of Applicants' Academic and School Indicators for the Most Competitive Tier (in SAT Context Condition)**

<b>Academic or School Indicator</b>	<b>Applicant</b>	
	<b>Low-SES</b>	<b>High-SES</b>
Unweighted HSGPA	3.90	3.85
Weighted HSGPA	4.13	4.27
Number of honors/AP Classes	10	15
SAT critical reading   HS median   Zip code median	670   430   440	720   600   590
SAT math   HS median   Zip code median	660   450   460	720   600   610
SAT writing   HS median   Zip code median	640   440   450	700   590   600
High school name	Smallville High	Suburbia High
State	New Jersey	Pennsylvania
Institutional control	Public	Public
Number of students	362	827
Graduation rate	63%	97%
Parent 1 education level	High school diploma	Doctorate degree
Parent 2 education level	Some high school	Master's degree

Note: AP = Advanced Placement; SES = socioeconomic status

Note: Horizontal bars distinguish between academic and high school information as well as parental education, which appears in all conditions. The high school median and zip code median SAT scores only appear in the SAT context condition. Weighted HSGPA was computed by adding ½ point for an honor class and a full point for an AP class; a description of the weighting approach was provided to participants within the simulated application files.

**Table 2:**  
**Summary Statistics by SAT Context and Control (Mean Values and Standard Deviations)**

	<b>No SAT context</b>	<b>SAT context</b>
Selectivity Tier 2	0.34 (0.48)	0.34 (0.47)
Selectivity Tier 3	0.37 (0.47)	0.39 (0.49)
Female	0.54 (0.50)	0.60 (0.49)
Person of color	0.21 (0.41)	0.27 (0.45)
Parental education	6.19 (1.96)	6.40 (1.81)
Admission experience	4.62 (1.53)	4.40 (1.45)
Work at alma mater	0.43 (0.50)	0.39 (0.49)
Grades are most important to merit criterion	0.46 (0.50)	0.46 (0.50)
Test scores are most important to merit criterion	0.09 (0.28)	0.12 (0.32)
Submitting test scores is optional	0.17 (0.37)	0.25 (0.43)
Consider scores for low-income apps differently	0.67 (0.47)	0.62 (0.49)
High school info is similar across applications	0.50 (0.50)	0.64* (0.48)
Quality of high school info received	4.26 (0.74)	4.36 (0.77)
Write paragraph explaining decision	0.48 (0.50)	0.43 (0.50)
Demonstrated interest considered in admission	0.43 (0.50)	0.56* (0.50)
Institutional fit considered in admission	0.72 (0.45)	0.76 (0.43)
Non-cognitive attributes considered in admission	2.51 (1.02)	2.51 (1.06)

Note: Chi-square and one-way ANOVA tests were conducted on all covariates to examine differences in mean values across treatment and control groups. ANOVA = analysis of variance.

\* $p < .05$



**Table 3:**  
**Logistic Regression results and AMEs for Admission Recommendations ( $n = 321$ )**

<b>Variable</b>	<b>Low-SES, low- grit essay</b>	<b>Low-SES, high- grit essay</b>	<b>High-SES, low- grit essay</b>	<b>High-SES, high-grit essay</b>
SAT context	1.53 (0.41)	1.61† (0.43)	1.76† (0.47)	1.33 (0.36)
AME	0.08	0.09†	0.11*	0.05

Note: Acceptance recommendations were coded as 0 = deny or wait list, 1 = accept. Coefficients presented are odds ratios; standard errors are in parentheses. AME = average marginal effects; SES = socioeconomic status.

† $p < .10$ . \* $p < .05$ .

**Table 4:**  
**Results for Multiple Regression Analyses Predicting Ratings of Academic Record,  
 Extracurricular Activities, and Personal Statements in SAT Context Condition ( $n =$   
 321)**

<b>Variable</b>	<b>Low-SES, low- grit essay</b>	<b>Low-SES, high- grit essay</b>	<b>High-SES, low- grit essay</b>	<b>High-SES, high-grit essay</b>
Academic record	0.08 (0.08)	0.17* (0.08)	0.14† (0.08)	0.10 (0.07)
Extracurricular activity	0.10 (0.09)	0.11 (0.10)	0.14 (0.09)	0.16† (0.08)
Essay	0.06 (0.11)	0.15 (0.12)	-0.05 (0.12)	0.25* (0.11)

Note: Academic record, extracurricular activities, and personal statements were all coded on a 6-point scale (1 = *very poor* to 6 = *excellent*). Standard errors are in parentheses. SES = socioeconomic status.

† $p < .10$ . \* $p < .05$ .

**Table 5:**  
**Results for Multilevel Logistic Regression Analyses Predicting Acceptance**  
**Recommendations (Level 1) Nested Within Participants (Level 2) (*n* = 321)**

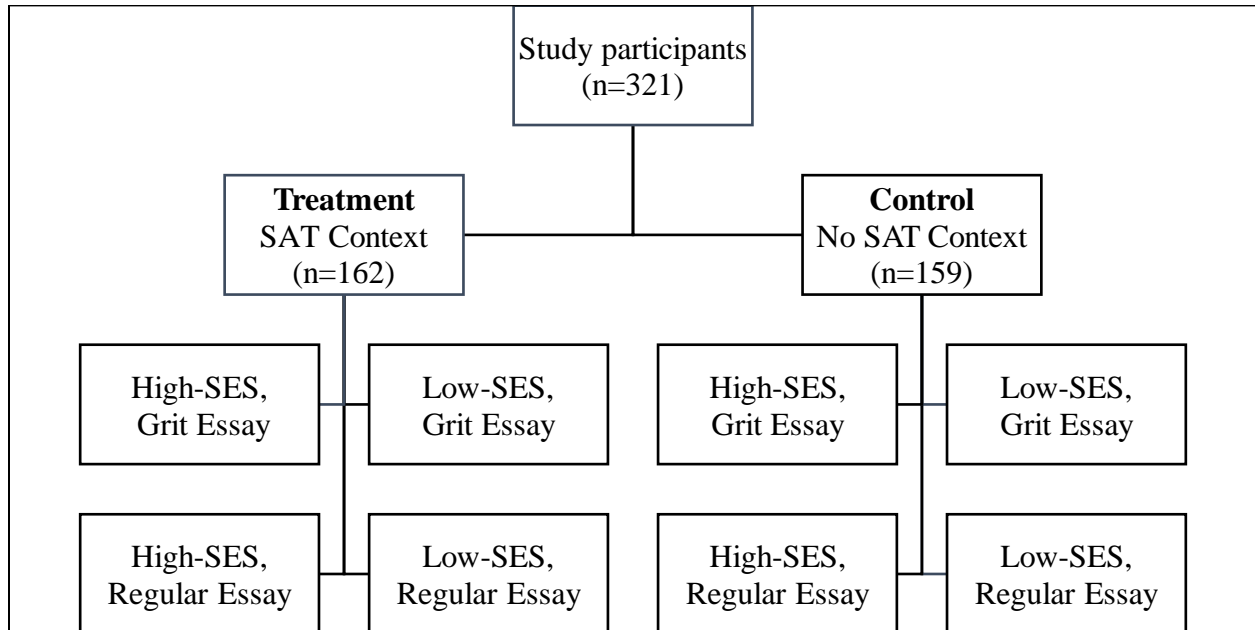
<b>Independent Variable</b>	<b>Model 1</b>	<b>Model 2</b>
SAT context	2.44* (0.94)	2.92* (1.35)
Low-SES applicant	1.06 (0.19)	1.13 (0.27)
High-grit applicant	1.06 (0.19)	1.20 (0.29)
Selectivity Tier 2	2.22 (1.21)	2.22 (1.21)
Selectivity Tier 3	7.29** (4.23)	7.31** (4.24)
Interactions		
SAT Context x low-SES applicant		0.89 (0.31)
SAT Context x high-grit applicant		0.78 (0.28)
Participant attributes		
Female	0.84 (0.33)	0.84 (0.33)
Person of color	0.42† (0.19)	0.42† (0.10)
Parental education	0.94 (0.10)	0.94 (0.10)
Admission experience	0.82 (0.12)	0.82 (0.11)
Work at alma mater	1.18 (0.46)	1.18 (0.46)
Admission office attributes		
Grades are most important merit criterion	0.78 (0.33)	0.78 (0.33)
Test scores are most important merit criterion	1.09 (0.75)	1.09 (0.75)
Submitting test scores is optional	3.39* (1.82)	3.39 (1.82)
Consider scores for low-income apps differently	1.18 (0.51)	1.18 (0.51)
High school info is similar across applications	1.12 (0.43)	1.12 (0.43)
Quality of high school info received	0.65† (0.17)	0.65† (0.17)
Write paragraph explaining decision	0.38* (0.17)	0.38* (0.17)

Demonstrated interest considered in admission	2.79*	2.79*
	(1.15)	(1.15)
Institutional fit considered in admission	0.35*	0.35*
	(0.17)	(0.17)
Non-cognitive attributes considered in admission	1.14	1.14
	(0.23)	(0.23)

Note: Acceptance recommendations were coded as 0 = deny or wait list, 1 = accept. Coefficients presented as odds ratios; standard errors are in parentheses. Reference groups were tier 1 for selectivity and curriculum rigor as the most important merit criterion. SES = socioeconomic status.

† $p < .10$ . \* $p < .05$ . \*\* $p < .001$ .

**Figure 1:  
Experimental Randomization**



Note: SES = socioeconomic status