



Commentary: The Questions We Should Be Asking About Socially Responsible College Admission Testing

Citation

Klugman, E.M., An, L., Himmelsbach, Z., Litschwartz, S.L. and Nicola, T.P. (2021), Commentary: The Questions We Should Be Asking About Socially Responsible College Admission Testing. Educational Measurement: Issues and Practice. <https://doi.org/10.1111/emip.12449>

Permanent link

<https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37368826>

Terms of Use

This article was downloaded from Harvard University's DASH repository, WARNING: No applicable access license found.

Share Your Story

The Harvard community has made this article openly available. Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

This is the submitted version of the following article:

Klugman, E.M., An, L., Himmelsbach, Z., Litschwartz, S.L. and Nicola, T.P. (2021),
Commentary: The Questions We Should Be Asking About Socially Responsible College
Admission Testing. Educational Measurement: Issues and Practice.

which has been published in final form at: <https://doi.org/10.1111/emip.12449>

*This article may be used for non-commercial purposes in accordance with Wiley Terms and
Conditions for Use of Self-Archived Versions.*

The Questions We *Should* Be Asking About Socially Responsible College Admission Testing

Emma M. Klugman, Lily An, Zachary Himmelsbach, Sophie L. Litschwartz, and Tara P. Nicola

Harvard Graduate School of Education

Author Note

Correspondence concerning this article should be addressed to Emma M. Klugman,
Harvard Graduate School of Education, 14 Appian Way, Cambridge, MA 02138. Email:
eklugman@g.harvard.edu

Abstract

While we agree with Koljatic et al. (2021) that the college admission testing industry can and should ensure their products promote educational equity, we do not believe the corporate social responsibility framework for which the authors advocate will assist testing agencies in achieving that aim. Koljatic et al.'s argument for corporate social responsibility in testing is rooted in flawed logic, and the notion of corporate social responsibility itself has unclear and undesirable implications for the testing industry. In this commentary, we offer four examples of the types of questions that test developers, researchers, and policymakers seeking socially responsible uses of admission testing should consider instead. These questions address concerns about test validity, the added value of test scores in the admission process, the burdens tests place on marginalized students, and the perpetuation of historic biases.

Keywords: college admission, high-stakes tests, corporate social responsibility, testing industry

The Questions We *Should* Be Asking About Socially Responsible College Admission Testing

The issue of what role testing agencies should play in facilitating the socially responsible use of their college admission tests is an important one. An increasingly hostile climate towards standardized testing in the undergraduate admission process (Carnevale et al., 2020), coupled with a burgeoning test-optional movement among the nation's selective universities (FairTest, 2021b), has called into question not only the purpose of these assessments but also their fairness and implications for reproducing inequities in U.S. higher education.

In their article, "College Admission Tests and Social Responsibility," Koljatic et al. (2021) argue that the responsibility to address social criticism and rejection of admission testing, as well as to ameliorate the social ills that tests promulgate, falls largely on the shoulders of testing agencies. They claim that these companies have avoided embracing such responsibility. Furthermore, they assert that having testing agencies embrace an ethic of corporate social responsibility (CSR) offers a panacea for these issues. While we agree that testing agencies can and should take responsibility for ensuring the socially responsible use of their products, we believe neither that testing companies have ignored questions of equity with their college admission tests, nor that advocating for agencies to embrace corporate social responsibility will bring about needed change.

In this commentary, we present a rebuttal to Koljatic et al. (2021)'s argument that CSR among testing agencies is a productive means of addressing educational inequities. Specifically, we assert both that Koljatic et al. (2021)'s argument applying Nike's adoption of CSR to the testing industry is flawed and that the framework of CSR has unclear and undesirable applications for the testing industry. We conclude by presenting examples of four questions that

test users, researchers, policymakers, and test developers wishing to advance the cause of socially responsible admission testing should consider instead. These questions address concerns about test validity, the added value of test scores in the admission process, the burdens tests place on marginalized students, and the perpetuation of historic biases.

Why Corporate Social Responsibility Isn't Quite Enough

We primarily take issue with two of Koljatic et al.'s arguments (2021): that CSR transformed Nike and ameliorated many of the ethical issues the company was facing; and that therefore CSR is helpful for guiding testing companies as they wade through ethical concerns with the use of their college admission tests.

First, Koljatic et al.'s (2021) assertion that corporate social responsibility has transformed Nike is inaccurate. Yes, Phil Knight changed the company's public stance on its responsibility for factory wages and working conditions in the late-1990s. But how effective has Nike leadership been at enforcing its labor standards since then? The evidence for transformative change is mixed at best, as scandals continue to hit the news regularly—see for instance a 2020 *Washington Post* report suggesting that Uighurs have been forced to work in some Nike factories (Fifield, 2020). What *has* changed is how Nike responds to such reports. Their public positions now clearly state that they accept addressing such abuses as part of their responsibility. (e.g., Nike, 2020). However, it is unclear whether the CSR practices Nike adopted have prevented abuses or if the resulting approach is essentially a kind of whack-a-mole game in which Nike leadership stomps out scandals as they pop up.

Nike's challenge, as described by Koljatic et al. (2021), is that the harmful practices documented in their factories were not implemented by Nike itself but by its contractors. Similarly, the most harmful uses of test scores may not be perpetrated by the testing companies

but by the colleges using them for admission purposes or by schools and test preparation companies adopting unsavory practices (Jump, 2021; Koretz, 2017). However, the analogy between Nike and testing companies breaks down since the colleges are not contractors for testing companies. This key difference calls into question the extent to which testing companies have control over the use of scores in the same way that Nike has over its production processes. The fix suggested by the authors—embracing an ethic of corporate social responsibility that emphasizes transparency and accountability—may stunt the bad press the testing companies find themselves in but fail to improve equity in the college admission process, which seems to be the ultimate goal.

The authors' framing of the analogy has other flaws. For example, whereas Nike faced significant pushback from consumers actively boycotting their products, the main consumers of admission tests—the colleges and universities that use them—are not applying that same level of pressure on testing agencies. Nike's customers called on the company to change, but higher education institutions are not demanding testing agencies dramatically re-envision their products and how they are used. Neither are higher education institutions boycotting the tests. Due to test center closures and other difficulties facing students during the COVID-19 pandemic, many institutions went test-optional (FairTest, 2021b). Notably, however, few have adopted a fully "test-blind" admission policy or decided to institute their own admission tests (FairTest, 2021a). Testing agencies are not incentivized to change if their primary consumers continue to rely on tests as part of their admission processes: as long as colleges accept test scores, students will continue to take the tests. In sum, Nike as a case study does not offer strong evidence of the positive effects of CSR in addressing complex ethical issues nor can we easily compare Nike with testing organizations.

Even if we put aside issues with the Nike analogy, we believe that CSR is not a helpful concept to guide the testing industry in navigating its role in promoting educational equity.

CSR for Testing Companies: Ill-defined and Undesirable

Based on the definition of CSR given in Koljatic et al. (2021), the authors hold testing companies responsible for disparities in test scores, maintaining that score differences cannot be blamed solely on existing societal inequalities. We agree that unequal test scores can exacerbate societal inequality, but we do not think that the focus on test scores themselves leads to reasonable solutions that testing companies should enact. If test score differences cannot be blamed solely on “inequalities in our education system” but themselves “perpetuate” such inequalities, then it seems the authors would prefer that test scores were equal across groups.

How could equal test scores across groups be achieved by testing companies? If admission tests were simply predictive algorithms, then score disparities between groups could be artificially erased. This would require that the relationship between correct responses to test items and final test scores be altered by some algorithmic boost or penalty to the scores of different groups.

However, test scores have to be calculated based on test performance if they are to have any meaning. Test items ideally assess some construct-relevant concept and have correct and incorrect answers (or higher and lower scoring answers). Candidates who correctly respond to more test items must, then, receive higher test scores. A CSR framing of testing companies’ responsibilities for undesirable test score disparities points to a solution that we do not think testing companies can responsibly impose. Koljatic et al. (2021) may have stopped short of directly advocating for score adjustments because they know that such a practice would be considered unfair, not to mention that any more nuanced score adjustments that attempt to take

into account the intersectionalities between race, class, income, and educational resources could never be sufficiently captured by any algorithm.

Or perhaps the authors do not intend for testing companies to equalize test scores as we just described, but believe instead that CSR requires testing companies to adopt other strategies to bring about more equitable rates in college enrollment, particularly at selective institutions. One such strategy would be providing colleges with measures of students' "opportunity to learn" to contextualize test scores. While we think that such efforts have potential, we have already seen widespread public criticism in response to the College Board's introduction and rapid withdrawal of its "adversity scores" in 2019 (Jaschik, 2019; Hartocollis, 2019), and again, Koljatic et al. (2021) stop short of advancing such a specific solution. In summary, the notion of CSR for testing companies implies solutions which would compromise the fundamental relationship between test taker performance and test scores, and would also likely meet with animosity.

The Questions We *Should* Be Asking About Socially Responsible College Admission Testing

Given the complexity of enacting a CSR framework for college admission testing, what then is a viable means for ensuring testing agencies promote socially responsible uses of their tests? We believe that major testing stakeholders—namely, higher education institutions and policymakers—should demand that testing agencies produce more evidence defending tests and their place in the world of college admission. Without these stakeholders putting pressure on testing agencies, testing companies will not be incentivized to reimagine their role in how test scores are utilized or otherwise address the societal inequities their tests may perpetuate. We

suggest four examples of the types of broad questions that these stakeholders and testing companies should consider pursuing below.

1. Does the Determination of Test Validity Increase Inequality?

Both the ACT and the College Board claim their exams measure proficiency on high school material needed to succeed in college. In practice, exam validity is assessed by estimating how well these tests predict first year college GPA and retention into the second year (ACT, 2020; Westrick et. al., 2019). Using this metric of validity can be problematic, since a test that only measures a student's socioeconomic status also performs well by this metric. For example, Rothstein (2004) found that a substantial portion of the predictive power of the 1990s version of the SAT came from its ability to predict school-level demographics. However, replacing college admission tests with a measure of SES would *not* be widely understood as a valid measure of academic preparedness for college. Since the current validity procedure pushes college admission tests towards being measures of socioeconomic status even if that is not the test makers' intention, it is important for testing agencies to justify the validity of their tests—not just by how showing well they predict early college outcomes, but also by demonstrating that the content of the exam is substantively important for students to know before entering college. Furthermore, in assessing the predictive validity of admission exams, test makers should publish more research on how much of the tests' predictive validity is a function of serving as a proxy predictor of student background, and how much additional explanatory power they provide.

2. How Does Admission Testing Affect College-Going Students?

It is important to reflect on whether the benefits of the college admission testing system outweigh the costs to students, given that students generally have to take the exams to attend a four-year college. College Board researchers argue the SAT's benefits are that the tests identify

“college ready” students that might otherwise have been overlooked and that they can be used to target additional support to those that need it (Westrick et. al., 2019). Similarly, the ACT (2019) argues that its tests allow students to know if they are prepared for college. Justifying admission tests, though, requires additional research to prove that such benefits materialize. In fact, some research suggests the ACT’s argument is incorrect—that there is no such thing as college “overmatch” and that all students are best off attending a four-year college regardless of their test scores (Kozakowski, 2020; Smith, 2020). Therefore, testing companies could do more to study situations where the promises of these benefits are fulfilled. For example, the University of Michigan instituted a program that in part used SAT scores to find high-scoring, low-income students and increased college enrollment and persistence by guaranteeing them full scholarships (Dynarski et. al., 2021). Researching other such avenues that support and clarify these benefits to students would further an argument about the utility of admission tests.

On the other side of the ledger, testing companies should also look for ways to reduce the costs to students, both in terms of money and time. The financial and time costs of completing the ACT and SAT exams are significant burdens, especially if students are expected or encouraged to take the test multiple times (Mattern et al., 2018; Goodman et al., 2020).¹ Preparation is also a factor; the College Board (2019) recommends students spend 6 to 20 hours studying for the SAT. Current efforts to more closely align tests with the high school curriculum are one step in the correct direction by reducing the time burden of studying for admission exams (ACT, 2020; Westrick et. al., 2019). Another idea that requires further investigation could be to allow students in states that already mandate high school level accountability exams to use those scores for college admission purposes.

¹ The ACT costs \$80 to take and the SAT \$68. In addition, the ACT and SAT charge \$15 and \$12 for sending each score, respectively.

3. How Fair is College Admission Without Tests?

A common argument in favor of maintaining college admission test requirements is that these tests comprise the least “gameable” portion of an admission application. A typical college application consists of a student’s transcript, essays, recommendation letters, application form, and test scores. Realistically, cheating can affect most portions of a college application. For example, there may be overstated extracurriculars, paid-for essays, and distortions of GPA from tutoring or homework-related grade inflation. Even recommendation letters can be paid for or are dependent on a student having interactions with adults that may be influenced by race and class status. Though we do not presume that most college applications suffer from these sorts of inflationary pressures, this potential issue may become more salient if standardized test scores are removed from the application. A college admission test may also be tutored, but it is very rare that someone besides the student took the test, or that the student was able to access help during the test. Without test scores, the other components of a college application will necessarily exercise more influence on admission decisions, which can be less fair for students without the financial means or personal desire to “game” their application. Work done on the fairness of letters of recommendation, student extracurriculars, and other pieces of a college application could add to the argument of the value-add of standardized tests.

4. How Do Testing Agencies Create New Products When Using Biased Historical Data?

Testing companies are always innovating, creating new exams and enrollment management products. The reality, however, is that these products are informed by historical data, including testing data which reflects historical processes and biases. To advance more equitable behavior and outcomes in college admission, testing companies must seek to acknowledge, quantify, and mitigate the impact that historical biases have on the data they and

others use, and be careful not to design products or interventions that reinforce these. For example, if a testing company wishes to recommend potential colleges to students by modeling whether a student's profile is similar to the profiles of past admitted students who have succeeded at a given institution (that is, training some sort of admission algorithmic to mimic human decisions), then such an algorithm would learn to perpetuate past biases and admit mostly candidates from historically advantaged groups. Beyond test scores, every aspect of a college application can be colored by past biases, many of which are understudied. More research is needed to understand the biases inherent in data and identify ways of ensuring those negative impacts are not further reinforced.

Moving Forward

Despite our concerns with some of the arguments Koljatic et al.'s (2021) advance, we share the authors' belief that socially responsible use of admission tests is important and that testing companies have a role to play in promoting such use. Furthermore, those who study measurement and psychometrics cannot be passive observers if test scores are promoting educational inequities. However, unlike Koljatic et al. (2021), we believe that advocating for testing agencies to adopt CSR will not bring about the authors' desired outcomes. It is not realistic to expect companies will change unless they are pressured to do so by their stakeholders and customers. We presented four questions that exemplify those that admission test stakeholders should push testing agencies to answer. Companies' responses to these questions can then be used as a basis for advocating that they change their practices and priorities where necessary.

As current doctoral students in education policy who care about applied educational measurement, we do not speak for testing companies or college admission offices. We have

attempted to take the perspective of eager researchers in measurement who care about the field's continuity and ability to effect positive social change. We believe the educational measurement community—including testing agencies—has a responsibility to promote the socially responsible use of test scores and that it can and must do more.

References

ACT. (2020). *ACT technical manual* (Version 2020.1).

www.act.org/content/dam/act/unsecured/documents/ACT_Technical_Manual.pdf

Carnevale, A., Schmidt, P., & Strohl, J. (2020). *The merit myth : How our colleges favor the rich and divide America*. The New Press.

College Board. (2019, August 26). What is the best way to prepare for the SAT? *College Board Blog*. <https://blog.collegeboard.org/what-best-way-prepare-sat>

Dynarski, S., Libassi, C. J., Micheltore, K., & Owen, S. (2021). Closing the gap: The effect of reducing complexity and uncertainty in college pricing on the choices of low-income students. *American Economic Review*, 111(6), 1721-1756.

<https://doi.org/10.1257/aer.20200451>

FairTest. (2021). *69 Campuses with “test-blind,” “test free,” or “score free” admissions policies for fall 2021 (or beyond)*. FairTest.

<https://www.fairtest.org/sites/default/files/Test-Blind-Admissions-List.pdf>

FairTest. (2021, April 28). *1,400+ U.S. four-year colleges and universities will not require ACT/SAT scores for fall 2022 entry* [Press release]. <https://www.fairtest.org/1400-us-four-year-colleges-and-universities-will-no>

Fifield, A. (2020, February 19). China compels Uighurs to work in shoe factory that supplies Nike. *The Washington Post*. https://www.washingtonpost.com/world/asia_pacific/china-compels-uighurs-to-work-in-shoe-factory-that-supplies-nike/2020/02/28/ebddf5f4-57b2-11ea-8efd-0f904bdd8057_story.html

- Goodman, J., Gurantz, O., & Smith, J. (2020). Take two! SAT retaking and college enrollment gaps. *American Economic Journal: Economic Policy*, 12(2), 115-158.
<https://doi.org/10.1257/pol.20170503>
- Hartocollis, A. (2019). SAT 'adversity score' is abandoned in wake of criticism. *The New York Times*. <https://www.nytimes.com/2019/08/27/us/sat-adversity-score-college-board.html>
- Jaschik, S. (2019). New SAT score: adversity. *Inside Higher Ed*.
<https://www.insidehighered.com/admissions/article/2019/05/20/college-board-will-add-adversity-score-everyone-taking-sat>
- Jump, J. (2021). Ethical College Admissions: Detrimental. *Inside Higher Ed*.
<https://www.insidehighered.com/admissions/views/2021/03/29/use-test-scores-admissions-disadvantages-some-students-and-some-colleges>
- Koljatic, M., Silva, M., & Sireci, S. G. (2021). College admission tests and social responsibility. *Educational Measurement: Issues and Practice*. <https://doi.org/10.1111/emip.12425>
- Koretz, D. (2017). *The testing charade: Pretending to make schools better*. The University of Chicago Press.
- Kozakowski, W.C. (2020). *Essays on Higher Education and Inequality* [Unpublished doctoral dissertation]. Harvard University.
- Mattern, K., Radunzel, J., Bertling, M., & Ho, A. D. (2018). How should colleges treat multiple admissions test scores?. *Educational Measurement: Issues and Practice*, 37(3), 11-23.
<https://doi.org/10.1111/emip.12199>
- Nike. (2020, November 17). *NIKE, Inc. Statement on Forced Labor, Human Trafficking and Modern Slavery for Fiscal Year 2020* [Press release]. <https://purpose.nike.com/nike-statement-on-forced-labor>

- Rothstein, J. M. (2004). College performance predictions and the SAT. *Journal of Econometrics*, 121(1-2), 297-317. <https://doi.org/10.1016/j.jeconom.2003.10.003>
- Smith, J., Goodman, J., & Hurwitz, M. (2020). *The economic impact of access to public four-year colleges* (No. w27177). National Bureau of Economic Research. <https://doi.org/10.3386/w27177>
- Westrick, P. A., Marini, J. P., Young, L., Ng, H., Shmueli, D., & Shaw, E. J. (2019). *Validity of the SAT® for predicting first-Year grades and retention to the second year*. College Board. <https://collegereadiness.collegeboard.org/pdf/national-sat-validity-study.pdf>