

EDUCATION



*Untitled*, Grace Harris,  
IUE Student 2024

# SCRUTINIZING THE STANDARDS: A LITERATURE REVIEW OF THE ADVANTAGES AND DISADVANTAGES OF STANDARDIZED TESTING

**Ella Carbrey**

## **Abstract**

For schoolchildren in the twenty-first century, the weeks of standardized testing each year mean more than just hours of tedium and boredom; for some, they mean the difference between moving to the next grade with their peers or being held back, or whether they get placed in a gifted or ESL class. Those at the front of the room are not free from the scrutiny either as the assessment of teachers' job performance has become more reliant on the testing results of their students, and schools are even shut down when students are consistently poor performers. As the pressure surrounding these tests has risen, so has the research exploring whether this is a beneficial change for education. Jennings and Lauen (2016), Kaufman et al. (2015), Toldson and McGee (2014), Laurito et al. (2019), and Jacob and Rothstein (2016) all explore this same topic from different perspectives with their research. In the literature, though the perspectives of the articles differ, similarities emerge regarding the issue of achievement gaps in testing, the dangers of tests being so high stakes for students and educators, and the importance of informing professionals in academia regarding how to best interpret test scores.

*Key words: standardized testing, gifted, ESL class, yearly, poor performers, teachers' job performance, high stakes for students, high stakes for educators, interpret test scores, grade promotion, pressure, value*

For schoolchildren in the twenty-first century, the weeks of standardized testing each year mean more than just hours of tedium and boredom; for some, they mean the difference between moving to the next grade with their peers or being held back, or whether they are placed in a gifted or ESL class. Those at the front of the room are not free from the scrutiny either. The assessment of teachers' job performance has become more reliant on students' testing results, and schools are even shut down when students are consistently poor performers. As the pressure surrounding these tests has risen, so has the research exploring whether the high stakes testing of the 21st century helps or harms students. Jennings and Lauen (2016), Kaufman et al. (2015), Toldson and McGee (2014), Laurito et al. (2019), and Jacob and Rothstein (2016) all explore this same topic from different perspectives with their research. Though the perspective of each article differs, patterns emerge regarding the issue of achievement gaps in testing, the dangers of tests being too high stakes for students and educators, and the importance of informing professionals in academia regarding how to best interpret test scores. In the current article, the research, and findings from the above five articles will be explored. The varying perspectives of the issue and its causes, consequences, and solutions are described.

### **Issue**

The primary issue addressed by each of the authors was whether standardized tests are valid, and a large part of the concern with their validity is a result of the achievement gap found between students of different races. Jennings and Lauen (2016) focused their discussion on whether the accountability of schools based on testing scores helped or hurt academic achievement in those schools, and subsequently found that, "the black-white achievement gap has actually increased as a result of accountability" (p. 223). Similarly, Kaufman et al. (2015) looked at standardized testing through the lens of surveys asking

teachers and school leaders their opinions of high-stakes testing, finding that in schools where there are high rates of students receiving free or reduced-price lunch, "Significantly more teachers ... reported concerns about the difficulty of math and ELA tests" (p. 4). The focus of Toldson and McGee's (2014) argument centered around the achievement gap, specifically in the SAT and ACT, and how historically, "Black students' average scores have been the lowest among all racial groups" (p. 1). The topic of the achievement gap was similarly front and center in the work of Laurito et al. (2019) as they focused on the effects of crime in a student's neighborhood on their standardized test scores, observing that, "African American and Hispanic middle school students are more likely to report feeling unsafe in the classroom and on school grounds than white and Asian peers who attend the same schools" (p. 143). Finally, Jacob & Rothstein (2016) investigated standardized testing by warning other researchers of the dangers of the improper use of testing data, stating that test scores, "are generally not unbiased measures of student ability" (p. 86). A consensus reached by researchers was that the achievement gap is caused by a multitude of factors and is negatively affecting many students while undermining the validity of standardized tests themselves.

While most sources touched on the achievement gap, others also discussed the validity of the tests in measuring their intended subject. In their discussion of the racial achievement gap in ACT and SAT scores, Toldson and McGee (2014) pointed out that these tests are not only harmful to students of color but have also not been shown to be accurate predictors of college GPA, as displayed in research shared by the National Association for College Admission Counseling (Hiss & Franks, 2014, as cited by Toldson & McGee, 2014). Jacob and Rothstein (2016) warned researchers in other fields of this questionable validity, explaining that, "[researchers] frequently misuse test score measures, with potentially serious consequences for their analyses" (p. 87). There are questions in academia regarding whether standardized tests are

successful metrics of what they are intended to be quantifying.

### **Causes**

Both Jennings and Lauen (2016) and Kaufman et al. (2015) discussed the history of standardized testing in the US and how it has evolved, leading to the issues seen today. Jennings and Lauen (2016) discussed the history of accountability, and how the government began to assess schools based on the scores of their students on standardized tests, as well as how educators are assessed, writing that, “teacher evaluations were increasingly tied to these scores as well” (p. 220). Kaufman et al. (2015) discussed these developments as a background to their data from surveys of educators, explaining that states have choices regarding which tests their students will take, and these choices are affected by, “political pressures, including concerns about new standards and assessments constituting federal overreach and a ‘national curriculum’” (p. 2). Through data cited later in their article, Kaufman et al. (2015) showed that the testing group that a state was a part of affected how concerned their teachers were over the tests (Kaufman et al., 2015). Both sources discussed the progression of state and federal government towards creating these high-stakes tests, contributing to a multitude of issues.

While the above sources viewed standardized testing from a policy perspective, Toldson and McGee (2014) and Jacob and Rothstein (2016) discussed the tests themselves, and how their potential lack of validity could be the cause of the concerns. Toldson and McGee (2014) discussed this through the lens of the achievement gap between students of color and their white counterparts, and how some believe the lack of test validity is to blame for this phenomenon (Toldson & McGee, 2014). Jacob and Rothstein (2016) agreed that test validity should be questioned, warning that these tests, “[Measure] student ability with error” (p. 94). If high stakes tests are not accurate measures of the intended subjects, problems such as the ones outlined above

may result.

Laurito et al. (2019) inspected the issue more closely by examining the causes of poor test performance within individual schools (Laurito et al., 2019). Their research focused on the fact that test scores were worse after exposure to violence, showing that these scores are variable, but also that it is possible to insulate students from the effects of these events. Findings in this source also showed that, "Black students are overrepresented in schools with weak climates" (Laurito et al., 2019). This statistic could very well be one of the contributors to the testing achievement gap (Laurito et al., 2019).

### **Consequences**

One major consequence of standardized testing discussed is the idea of test preparation. Jennings and Lauen (2016) investigated how teachers often "teach to the test" because of federal accountability, in an effort to improve students' state test scores and help both students and teachers receive positive feedback after test results (Jennings & Lauen, 2016). Toldson and McGee's (2014) research focused on the ACT and SAT, and their discussion involved the trend for families of a higher socioeconomic class to pay for test preparation materials and tutors for their children (Toldson & McGee, 2014). Both articles discussed the result of having tests that are high stakes, either within a school or when trying to reach the next level of education, and how the focus on succeeding in these tests may undermine the actual knowledge of students (Toldson & McGee, 2014; Jennings & Lauen, 2016).

While other sources take note of how teachers' lesson plans react to testing, Kaufman et al. (2015) examined how educators felt about testing, which also relates to the concept of school environment that Laurito et al. (2019) inspected. Kaufman and colleagues (2015) discovered the pressure educators felt to help their students perform, finding that, "...[teachers] report that the work they do to prepare students for the PARCC [Partnership for Assessment of Readiness for

College and Career] test will take time away from other important work” (Kaufman et al., 2015, p. 3). On the other hand, Laurito et al. (2019) found that whether a student perceives their school’s environment to have a strong sense of community and safety affects their tests scores (Laurito et al., 2019). These researchers found that a consequence of high-stakes testing is educator stress, but if educators push students to the point of no longer feeling supported or safe, this could hurt more than help them (Kaufman et al., 2015; Laurito et al., 2019).

### **Solutions**

The idea of communication as a key part of the solution was a recurring theme in the research, with Kaufman and colleagues (2015) and Jacob and Rothstein (2016) both discuss the need for clear communication regarding how to interpret test scores and how they will be used to make decisions. Kaufman et al. (2015) discussed this in relation to the need to alleviate educators’ stress around high-stakes testing, proposing that educators should better understand how to interpret their students’ results, and that, “States and districts should also clearly communicate how—if at all—test results will be used to make teacher- and school-performance decisions” (Kaufman et al., 2015, p. 5). Jacob and Rothstein (2016) focused on the need for communication and education regarding how these test scores should be utilized, ultimately stating that, “There are no simple fixes, other than to be cautious in interpreting results” (Jacob & Rothstein, 2016, p. 104). There is consensus among researchers regarding the need for caution when interpreting test scores and using them to make important decisions or further research assertions (Kaufman et al., 2015; Jacob & Rothstein, 2016).

Other sources focused on how to solve the problem through the mindset of students. Toldson and McGee (2014) warned Black students of the harm of the achievement gap, advising that, “due to concerns about test validity concerning college entrance exams, the

Black community should not entertain social commentary that links low test scores to any functional impairment of the race” (Toldson & McGee, 2014, p. 2). Laurito and colleagues (2019) argued more broadly that the focus of schools should be creating a positive and safe environment for students to allow them to not only do their best on testing, but also live more successful academic lives (Laurito et al., 2019). Both sources had a student-centric proposed solution (Toldson & McGee, 2014; Laurito et al., 2019).

### **Conclusion**

These sources presented a wide variety of perspectives and beliefs; Laurito et al. (2019) inspected how the environment of a school could help or hurt the students in many ways, Toldson and McGee (2014) warned students of color of the questionable validity of college entrance exams, Jacob and Rothstein (2016) advised to other researchers the need for caution when inspecting test scores, Kaufman et al. (2015) displayed the concerns of educators surrounding standardized testing, and Jennings and Lauen (2016) described how test scores and teachers reacted to federal accountability. Though not all proposed clear and exact solutions, it is apparent that the implementation of further communication among policymakers, educators, students, and parents would be beneficial. ■



## REFERENCES

- Jacob, B., & Rothstein, J. (2016). The measurement of student ability in modern assessment systems. *The Journal of Economic Perspectives*, 30(3), 85–107. <http://www.jstor.org/stable/43855702>
- Jennings, Jennifer L. & Lauen, Douglas L. (2016). Accountability, inequality, and achievement: The effects of the no child left behind act on multiple measures of student learning. *RSF: The Russell Sage Foundation Journal of the Social Sciences*, 2(5), 220–241. <https://doi.org/10.7758/rsf.2016.2.5.11>
- Kaufman, J., Hamilton, L., Stecher, B., Naftel, S., Robbins, M., Garber, C., Ogletree, C., Faxon-Mills, S., & Opfer, D. (2015). What are teachers' and school leaders' major concerns about new K–12 state tests?: Findings from the American Teacher and American School Leader Panels. In What Are Teachers' and School Leaders' Major Concerns About New K–12 State Tests?: Findings from the American Teacher and American School Leader Panels, 1–8. *RAND Corporation*. <http://www.jstor.org/stable/10.7249/j.ctt19w72c7.1>
- Laurito, A., Lacoë, J., Schwartz, A. E., Sharkey, P., & Ellen, I. G. (2019). School climate and the impact of neighborhood crime on test scores. *RSF: The Russell Sage Foundation Journal of the Social Sciences*, 5(2), 141–166. <https://doi.org/10.7758/rsf.2019.5.2.08>
- Toldson, Ivory A. & McGee, Tyne. (2014). What the ACT and SAT mean for Black students' higher education prospects (Editor's Commentary). *The Journal of Negro Education*, 83(1), 1–4. <https://doi.org/10.7709/jnegroeducation.83.1.0001>