

Impact of Superscoring on Subgroup Differences

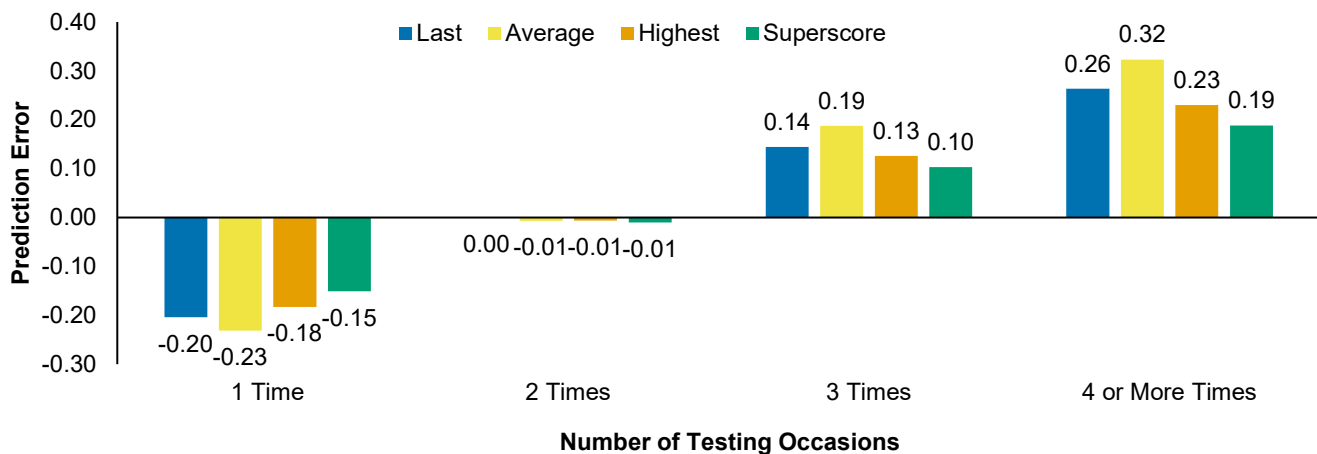
Krista Mattern, PhD, and Justine Radunzel, PhD

When applicants take the ACT® more than once, how do colleges and universities reconcile and make sense of the multiple scores? In terms of validity, fairness, and impact on subgroup differences, are certain score-use policies better than others? The focus of this issue brief is to summarize evidence on the validity and fairness of various score-use policies with an emphasis on superscoring. The intent is to arm higher education professionals with the most recent evidence to help support informed decision making on their own campus.

Validity and Fairness of Superscoring

ACT has been examining the validity and fairness of different scoring practices over the last several years.¹ Contrary to expectations, the results showed that scores based on the superscoring method (referred to as superscores) were just as predictive (actually, slightly more predictive) of first-year grades as compared to other scoring methods (recent, average, highest). Moreover, superscoring resulted in the least amount of differential prediction by the number of times a student tests. Interestingly, we found that first-year grades for students who tested more often were underpredicted even when prediction models were based on superscores (see Figure 1).²

Figure 1. Magnitude of Differential Prediction by Number of Testing Occasions and Four Composite Scoring Methods when ACT Composite Score is Held Constant at the Sample Mean of 23



Note: Prediction error is calculated by subtracting one’s expected FYGPA based on the overall model from the expected value based on the model that includes retesting subgroup indicators and the interaction between the ACT Composite score and retesting indicators. Negative values indicate overprediction; positive values indicate underprediction.

As shown in Figure 1, retesters performed better in college than what was expected based on their test scores. And this prediction error was minimized when superscores were used, as compared to the other scoring methods.



ACT.org/research

Current Study

The focus of the current study is to extend the previous research with an emphasis on further exploring the diversity implications of superscoring. Previous research indicates underserved students are less likely to retest as compared to their affluent peers.³ For students who test only once, superscoring has no impact on their ACT Composite score. Only students who retest have the potential to increase their ACT Composite score through superscoring, and the magnitude of this difference should be related to the number of times the student retests, in general. With that in mind, one potential concern or unintended consequence of superscoring is that subgroup differences will be exacerbated under this scoring policy. The focus of the current study is to investigate the extent to which superscoring increases, decreases, or has no impact on subgroup differences.

Using data on the 2018 ACT-tested graduating class, we compared the average ACT Composite score for various student subgroups (gender, race/ethnicity, family income, parental education level) based on their most recent ACT Composite score as well as a superscore ACT Composite score.

The results indicate that superscoring increased subgroup differences only marginally. On average, unstandardized differences are 0.17 larger (on a 1 to 36 scale) for superscores as compared to the most recent scores. For example, the average ACT Composite score for African American students is 5.3 points lower than White students (16.9 versus 22.2) when based on the most recent ACT Composite score. Comparatively, the average ACT Composite score for African American students is 5.4 points lower than White students (17.4 versus 22.8) when based on a superscore ACT Composite score, resulting in a difference in unstandardized differences of 0.10 (5.4 minus 5.3). When the results are disaggregated by the number of testing occasions, we see even smaller increases attributed to superscoring and often the results reverse where superscoring results in smaller subgroup differences.

In sum, the results indicate that subgroup differences are largely unaffected by the two scoring policies examined—most recent versus superscoring. Given that students tend to improve their scores through retesting and the high reliability of ACT scores, it is not surprising that results based on a student's most recent test record are quite similar to those based on superscoring. Also contributing to the finding of small to no differences based on superscoring is the relatively low frequency of retesting overall and retesting more than once. The results also suggest that the slight increases in unstandardized differences can be attributed to differences in retest rates among subgroups. If we improve retesting rates through programs and initiatives among groups who are less likely to retest, such as underserved students, these results suggest that superscoring may help reduce subgroup differences. Despite these positive findings, the results may change if retesting behavior changes significantly in the future in terms of who retests and how often.

Notes

1. 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.
2. Research on the SAT found similar findings: Boldt, R. F., Centra, J. A., & Courtney, R. G. (1986). *The validity of various methods of treating multiple SAT® scores* (College Board Report No. RR-86-4). New York, NY: The College Board.
3. Harmston, M. & Crouse, J. (2016). *Multiple testers: What do we know about them?* Iowa City, IA: ACT.