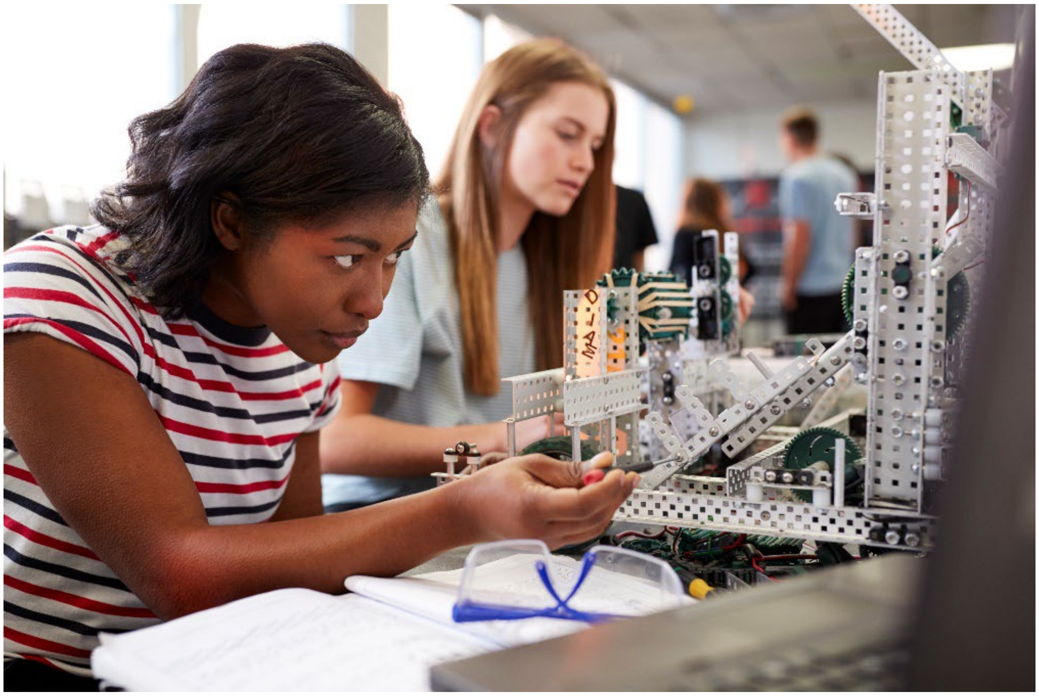
## Research Report

2024-05

# Equity in Education: An Examination of the Influences of Academic Preparation, Family Income, Race/Ethnicity, and Gender on ACT® STEM and ELA Scores

EDGAR I SANCHEZ, PHD



## Conclusions

This study investigates the allegations of socioeconomic, racial/ethnic, and gender biases on ACT® performance and evaluates the effectiveness of strategies implemented by ACT to address these biases. This study reveals that once academic readiness is accounted for, the disparities in ACT scores across different student subgroups are notably diminished. This paper also assesses ACT's efforts to promote equity, such as the inclusion of diverse perspectives in test development, fairness reviews, and the provision of resources like fee waivers and free test preparation. By examining the variance in ACT scores explained by students' academic backgrounds versus their socioeconomic status and demographics, this research contributes to the ongoing discussion on standardized testing fairness and underscores the importance of holistic evaluations of student capabilities.

## So What?

By highlighting the significant role of students' high school grades and coursework in determining ACT scores, the study underscores the need for equitable educational opportunities for all students. The findings challenge the perception of inherent biases in the ACT by demonstrating that observed subgroup differences can be substantially explained by academic factors, encouraging a more nuanced approach to interpreting standardized test scores and their fairness.

## Now What?

This study suggests that efforts to reduce disparities in standardized test scores should focus on addressing inequalities in academic preparation and access to advanced coursework, rather than modifying the test itself. This emphasizes educational reforms that ensure all students, regardless of socioeconomic status or background, have access to a high-quality education.

## About the Author

**Edgar I Sanchez**

Dr. Edgar I. Sanchez is a lead research scientist at ACT, where he studies postsecondary admissions, national testing programs, test preparation efficacy, and intervention effectiveness. Throughout his career, Dr. Sanchez's research has focused on the transition between high school and college and supporting the decision-making capacity of college administrators, students, and their families. His research has been widely cited in academic literature and by the media, including The Wall Street Journal, The Washington Post, USA Today, and the education trade press.

## Acknowledgements

The author wishes to thank Jeff Conway, Joyce Schnieders, and Dana Murano for their suggestions on previous drafts of this report.

# Introduction

The ACT® test, like many standardized assessments utilized in the American educational system, faces allegations of bias in its construction that may negatively affect certain groups of students who take the assessment. Three such biases frequently mentioned are socioeconomic bias, racial/ethnic and cultural bias, and gender bias.

When arguments are made about socioeconomic bias, it is often suggested that students from higher socioeconomic backgrounds are favored by the ACT and, as a result, receive higher scores. It is suggested that these students have more resources available to them, such as access to high-quality schools, private tutoring, and expensive test preparation courses. It is argued that these resources lead to better test performance, which then results in an advantage that creates a disparity between students from different backgrounds. For example, Kohn (2000) argued that socioeconomic status is a major source of variance in standardized test scores and suggested therefore that students' socioeconomic background significantly influences performance on standardized tests such as the ACT. Milner (2013) counters this argument by suggesting that standardized test scores will vary for students based in part on instruction and learning opportunities in addition to issues unrelated to education such as poverty, employment, and where homes are located.

Claims about racial and cultural bias largely pertain to inherent biases against certain racial/ethnic groups. This can be due to cultural references, language nuances, or questions framed in a manner that is more familiar to some groups than others. This argument states that this type of bias can adversely impact the performance of students from underrepresented racial and ethnic backgrounds. One example of this argument specifically directed at the ACT (FairTest, 2007b) suggested that the ACT did a better job at predicting outcomes for White students than it did for Black students.

The final bias that tends to be mentioned is gender bias. In this argument, it is said that some sections of the ACT may favor one gender over others, particularly in the areas of math and science, where there have been historically documented differences between gender groups. These differences are argued to represent gender bias in test scores. One such argument for gender bias in the ACT has been made by FairTest (2007a). In this article, the authors claimed that because female students tend to score lower on the ACT than male students, the test is necessarily biased against female students and will result in an underprediction of their ability.

At ACT, a number of methods have been utilized in order to prevent or minimize possible bias in The ACT test. For example, a diverse group of individuals are involved in the item development process, which helps to ensure that items are free from cultural, racial, and gender biases. Teams of experts from various backgrounds and specializations help identify and address potential biases in the questions before these questions are used in the ACT.

Additionally, ACT conducts external fairness reviews for all items prior to pretesting and then again for forms before they become operational. During the external content review, stimuli and items are evaluated for content accuracy as well as appropriateness of language and context. ACT invites external reviewers with knowledge and experience in relevant content areas, including high school teachers, to participate. Reviewers with various backgrounds are selected to ensure diversity in terms of gender, race, culture, and geography.

This development process at ACT also includes a comprehensive statistical review and validation process of test items. These reviews include statistical analysis to identify any items that show Differential Item Functioning (DIF) across different demographic groups. DIF is used to identify items that behave differently across different subgroups of students, and, if items are found to have DIF, they are either revised or removed.

Furthermore, to address socioeconomic considerations, the ACT offers fee waivers, free test preparation, and a vast network of test centers to help reduce the socioeconomic disparities in test access and preparation. For example, the fee waiver program enables qualifying students to take the ACT at no cost. These students are also granted access to free test preparation services to help them prepare to do their best on the ACT. Test centers are also selected in a strategic manner to ensure that students from lower socioeconomic backgrounds have equitable access to testing centers.

In addition to these efforts, ACT has also conducted research to explore the primary sources of variance in ACT scores. For example, McNeish et al. (2015) utilized a blockwise regression model with robust standard errors to analyze the relationship between cognitive and noncognitive traits and ACT scores. High school grade point average (HSGPA) was the main explanatory factor, explaining 20% to 31% of the variance in ACT scores. High school coursework contributed an extra 8% for reading and up to 17% for mathematics, while other high school characteristics (e.g., percent of school eligible for free or reduced-price lunch) explained 7% to 9% more of the variance in ACT scores. Socioeconomic and demographic factors had a smaller impact on ACT scores, explaining 4% or less of the variance after adjusting for other student and school characteristics. Importantly, the differences in average scores across different racial/ethnic groups, family income levels, and levels of parental education were significantly smaller after adjusting for HSGPA and high school coursework. The McNeish et al. (2015) study demonstrates the importance of academic preparation for performance both on the ACT and on postsecondary outcomes. This study, however, focused on subgroup differences in ACT Composite score.

What is lacking in the current literature is an understanding of how accounting for high school achievement may mitigate differences observed in ACT Science, Technology, Engineering, and Mathematics (STEM) and ACT English Language Arts (ELA) scores among students in different demographic groups (e.g., race/ethnicity, gender, family income). Since fall 2015, ACT has reported a STEM score, which is calculated as the average of the 1–36 mathematics and science scale scores rounded to the nearest integer (fractions of 0.5 or greater round up). Only students who receive scores on the mathematics and science tests receive an ACT STEM score. In fall 2015, ACT also began reporting a combined ELA score. The ACT ELA score is the rounded average of the English score, the reading score, and the 1–36 writing scale score. Only students who take all three of these tests can receive an ELA score. For the calculation of ELA scores, the sum of the writing domain scores is converted to a scale of 1–36. However, this 1–36 writing scale score is not reported independently.

The present study uses data from the ACT-tested graduating class of 2022 to explore the strength of the relationship between high school coursework, socioeconomic status, and student demographic characteristics with ACT STEM and ELA scores. This study endeavors to demonstrate that, after accounting for a student's high school academic achievement, socioeconomic status and demographics add little explanatory power, largely reducing the observed differences between student subgroups. In doing so, this study extends the current research by looking at subject-specific outcomes. This study explores the following three research questions:

1. What are the primary sources of variance observed in ACT STEM scores?

2. What are the primary sources of variance observed in ACT ELA scores?

3. Are subgroup differences in ACT STEM and ELA scores reduced after accounting for achievement and academic preparation?

# Method

## Analytical Sample

The present study uses data from the 2022 ACT-tested high school graduating class (N= 1,349,644). Due to the requirement of students having obtained a valid mathematics and science test score to receive an ACT STEM score and a valid English score, reading score, and writing score to receive an ACT ELA score, two independent samples from the 2022 graduating class are used for the present analysis. In the graduating class of 2022, 877,917 students qualified for inclusion in this study by having an ACT STEM score. Due to computational limitations (i.e., the full data set was too large to process with available computer resources) a random sample was utilized in this study. The ACT STEM sample consisted of 333,000 students (38%) and mirrored the percentages of students in the full sample by socioeconomic status and demographic characteristics (see Appendix for population and sample statistics). The ACT ELA sample consisted of 217,371 students who had valid ACT ELA scores. No computational issues arose when analyzing the ELA sample, and therefore random sampling a subset was not necessary as it was for the STEM sample. In both the ELA and STEM samples, about 50% of students identified as female, about 50% of students identified as White, about 45% to 46% of students did not report their family income, and both samples had similar English, math, social studies, and science GPAs (Table 1).

**Table 1.** Sample Demographic Characteristics

| Characteristic | | ELA | STEM |
|---|---|---|---|
| Gender | Female | 164,802 (50%) | 674,283 (50%) |
| | Male | 158,231 (47%) | 631,327 (47%) |
| | Other/Prefer not to respond/Missing | 10,158 (3%) | 44,021 (3%) |
| Race/Ethnicity | Asian | 17,435 (5%) | 54,464 (4%) |
| | Black | 29,922 (9%) | 153,579 (11%) |
| | Hispanic | 55,968 (17%) | 210,204 (16%) |
| | White | 164,103 (49%) | 708,950 (53%) |
| | Other | 24,039 (7%) | 78,019 (6%) |
| | Prefer not to respond/Missing | 41,724 (13%) | 144,415 (11%) |
| Family Income | < $36K | 38,170 (11%) | 146,282 (11%) |
| | $36K–$60K | 29,344 (9%) | 112,291 (8%) |
| | $60K–$100K | 40,263 (12%) | 156,038 (12%) |
| | > $100K | 75,009 (23%) | 318,624 (24%) |
| | Missing | 150,405 (45%) | 616,396 (46%) |
| GPA (mean (SD)) | English | 3.4 (0.67) | 3.4 (0.72) |
| | Math | 3.3 (0.73) | 3.3 (0.77) |
| | Social Studies | 3.5 (0.64) | 3.5 (0.68) |
| | Science | 3.4 (0.68) | 3.3 (0.72) |
| N | | 214,731 | 333,000 |

*Note*. Subject GPA ranges from 0.0 to 4.0.

## Measures

***ACT STEM and ACT ELA Scores.*** Official ACT STEM and ACT ELA scores were obtained from the 2022 graduating class cohort record. These ACT scores may have been attained during either school-day testing or a National test administration. For students who took the ACT more than once, the most recent score prior to graduating from high school was used in the study.

***High School Subject GPAs.*** Self-reported grades in up to 23 courses in English, mathematics, social studies, and natural science were averaged to calculate each student's subject GPA. The calculation of English GPA included grades in English 9, 10, and 11. The calculation of math GPA included grades in Algebra 1, Algebra 2, Geometry, other math beyond Algebra 2, and Computer Math. The calculation of science GPA included grades in Physical Science, Earth Science, General Science, Biology, Chemistry, and Physics. The calculation of social studies GPA included grades in U.S. History, American History, World History, World Civilizations, Government, Civics, Citizenship, Psychology, and other history classes.

Sanchez and Buddin (2015) demonstrated that students' self-reported subject GPA is highly correlated with students' subject transcript GPA. Based on their findings, students' course-specific grades were similar to transcript grades—the median exact agreement rate (i.e., selfreporting letter grades were the same as those in their transcripts) was 68%, and reporting within one letter grade ranged from 91% to 100%. They also noted that in English, mathematics, science, and social studies students tended to underreport their grades rather than overreport them. Other research also supports the use of self-reported GPA data for research purposes (Camara et al., 2003; Kuncel et al., 2005; Shaw & Mattern, 2009).

***Coursework Taken.*** High school course-taking patterns in English, mathematics, natural science, and social studies were considered for inclusion in the present study. In the case of English coursework, the vast majority of students had taken English 9, 10, and 11 at the time of test registration, which made it an unusable indicator. There was a similar situation for mathematics, where most students had taken at least Algebra 1, Algebra 2, and Geometry. There are other combinations of advanced mathematics beyond geometry. However, the combinations of Trigonometry, beginning Calculus, and other advanced math resulted in very low N counts. There is some meaningful variation between students who have taken only Biology; Biology and Chemistry; and Biology, Chemistry, and Physics. For social studies, there is no natural sequence of course taking. For these reasons, course-taking patterns were not included as an indicator of academic preparation.

***Taken Advanced Coursework.*** A self-reported indicator of having taken advanced coursework in English, mathematics, natural science, and social studies was used. This indicator included having taken AP, accelerated, and/or honors courses for each subject.

***Student Demographics.*** Self-reported student demographics included family income, gender, and race/ethnicity. These data were collected at the time of ACT registration. Students selected their estimated total combined parental income from nine options: less than $24,000, $24,000–$36,000, $36,000–$50,000, $50,000–$60,000, $60,000–$80,000, $80,000–$100,000, $100,000–$120,000, $120,000–$150,000, and more than $150,000. These categories were collapsed into four categories: less than $36,000, $36,000–$60,000, $60,000–$100,000, and more than $100,000. Students selected their self-identified gender from four options: male, female, another gender, and prefer not to respond. For the present analysis, the category another gender was combined with prefer not to respond and missing responses due to the low number of students selecting these options. Students self-identified their racial/ethnic background from seven options: American Indian/Alaska Native, Asian, Black/African American, Native Hawaiian/Other Pacific Islander, White, prefer not to respond, or none of these apply. This response in conjunction with self-identified Hispanic background were used to create six racial/ethnic categories: Asian, Black, Hispanic, White, Other, and prefer not to respond or missing response.

## Data Analysis

For both the ACT STEM and ACT ELA scores, the following linear model building was estimated (see Table 2). For each of the model blocks, estimates of the change in $R^2$ were calculated to evaluate the proportion of variance in each ACT score that was explained by each successive block of predictors. An overall $R^2$ is also reported for the full model, which includes all blocks. $R^2$ is a measure of the proportion of variance in the dependent variable (i.e., ACT STEM or ACT ELA scores) explained by the predictors in the model. $R^2$ is calculated using the formula $R^2 = 1 - \frac{Sum\ of\ squares_{residual}}{Sum\ of\ squares_{Total}}$; where the sum of squared residuals (also known as the sum of squared errors) represents the sum of the squared differences between the predicted values and the actual values of the dependent variable, and the sum of squares total represents the sum of the squared differences between the actual values of the dependent variable and its mean. It is expected that each successive block will add incrementally less change in $R^2$. Additionally, standardized parameters are presented.

**Table 2.** Model Block Design

| Block | STEM Model | ELA Model |
|---|---|---|
| **Block 1: High School Grades Earned** | Mathematics and Science GPA | English and Social Studies GPA |
| **Block 2: Advanced Coursework Taken** | Taken/Not taken AP, accelerated, or honors coursework in mathematics and science | Taken/Not taken AP, accelerated, or honors coursework in English and social studies |
| **Block 3: SES Demographics** | Family income | Family income |
| **Block 4: Gender and Race/Ethnicity Demographics** | Race/ethnicity<br>Gender | Race/ethnicity<br>Gender |

Marginal means were calculated for each of the demographic characteristics to evaluate if subgroup differences observed prior to adjusting for student achievement were reduced after accounting for academic preparation (i.e., HSGPA and advanced coursework taken) in the final model. The marginal mean was calculated as the average model-predicted ACT STEM or ACT ELA score for a given demographic characteristic (i.e., family income, race/ethnicity, and gender) when other continuous predictors were held at their mean and other categorical predictors are held at their proportion values. This can be interpreted as the model-estimated dependent value mean for a given demographic characteristic.

In the present analysis, cluster robust standard errors were utilized to account for student clustering in high schools. An alternative methodology would have been to utilize hierarchical linear modeling (HLM) to account for students nested within high schools. While HLM accounts for student clustering by implementing random intercepts and/or slopes for each school, cluster robust standard errors allow for the specification of correlated residuals within schools. An advantage of using cluster robust standard errors as opposed to HLM is being able to utilize all high schools, even those with a low number of students per high school (Clarke, 2008; Clarke & Wheaton, 2007; McNeish, 2014). Alternatively, HLM methodology requires a minimum number of students per high school to ensure stable estimates of the random components for each high school. Thus, utilizing cluster robust standard errors allows for the incorporation of schools with a low number of students per school while still providing stable estimates of the model.

# Results

## What are the primary sources of variance observed in ACT STEM scores?

Model 1 included both math and science GPA and accounted for 31.6% of the variance in ACT STEM scores. In this model, a change of one standard deviation in math GPA was associated with a 1.70 scale score increase in ACT STEM score, and a change of one standard deviation in science GPA was associated with a 1.22 scale score change in ACT STEM score (Table 3). Model 2, which added indicators for taking advanced coursework in math and natural science, accounted for 41.8% of the variance in ACT STEM scores. Taking advanced coursework in math was associated with a 2.83 scale score increase in ACT STEM scores, and taking advanced coursework in science was associated with a 1.45 scale score increase in ACT STEM scores, after accounting for GPAs in block 1. This model resulted in a 10.2% increase in $R^2$ over model 1. The $R^2$ associated with model 2 is taken as the total percentage of variance explained by grades and coursework taken combined.

**Table 3.** Blockwise STEM Regression Coefficients

| Predictor | | Model 1 Estimate | Model 2 Estimate | Model 3 Estimate | Model 4 Estimate |
|---|---|---|---|---|---|
| **Intercept** | | 20.95 | 19.18 | 18.24 | 19.56 |
| **Math GPA** | | 1.70 | 1.46 | 1.31 | 1.26 |
| **Science GPA** | | 1.22 | 0.90 | 0.78 | 0.80 |
| **Taken Advanced Coursework in Math** | | — | 2.83 | 2.68 | 2.55 |
| **Taken Advanced Coursework in Science** | | — | 1.45 | 1.34 | 1.44 |
| **Family Income** | < $36K | — | — | −0.88 | −0.55 |
| | $60K–$100K | — | — | 0.83 | 0.55 |
| | > $100K | — | — | 2.29 | 1.80 |
| | Missing | — | — | 1.25 | 1.03 |
| **Race/Ethnicity** | Asian | — | — | — | 1.97 |
| | Black | — | — | — | −2.35 |
| | Hispanic | — | — | — | −1.35 |
| | Other race/ethnicity | — | — | — | −0.45 |
| | Prefer not to respond/Missing | — | — | — | 0.75 |
| **Gender** | Female | — | — | — | −1.43 |
| | Another Gender/Prefer not to respond/Missing | — | — | — | 0.38 |
| **$R^2$** | | 31.6% | 41.8% | 45.8% | 50.5% |
| **$\Delta R^2$** | | — | 10.2% | 3.9% | 4.7% |

*Note*. Math GPA and Science GPA were standardized. All predictors entered their respective blocks with a significance of <.0001. They were significant at the <.0001 level in the final model.

Model 3 added family income to the model and accounted for 45.8% of total variance explained, which was a 3.9% increase in explained variance over model 2. In this model, students from family incomes of less than $36,000, $60,000–$100,000, greater than $100,000, and students who did not report family income had a 0.88 point lower, 0.83 point higher, 2.29 point higher, and 1.25 point higher ACT STEM score, respectively, than students whose family income was $36,000–$60,000, after controlling for the variables in previous blocks.[1] Model 4 added race/ethnicity and gender to the model. This model was associated with an explained variance of 50.5%, which was a 4.7% increase over model 3. In this model, Asian students, Black students, Hispanic students, other race/ethnicities, and students who preferred not to respond or did not respond to the race/ethnicity question had a 1.97 higher, 2.35 lower, 1.35 lower, 0.45 lower, and 0.75 higher ACT STEM scale score, respectively, than White students, after the other variables were held constant. Female students had a 1.43 lower ACT STEM score than male students, and students from another gender, those who preferred not to respond to the gender question, or those who did not respond to the gender question had a 0.38 higher ACT STEM score than male students, after the other variables were held constant.

In the blockwise regression models, we see that math and natural science GPA along with indicators for taking advanced coursework in math and natural science accounted for the largest percentage of variance in ACT STEM scores: 41.8%. The addition of demographic information such as family income, race/ethnicity, and gender only explained an additional 8.6% of the variance in ACT STEM scores.

## What are the primary sources of variance observed in ACT ELA scores?

Model 1, which included English and social studies GPA, accounted for the largest percentage of variance explained: 32% (see Table 4). In this model, an increase of one standard deviation in English GPA was associated with an increase of 2.26 scale score in ACT ELA score. An increase of one standard deviation in social studies GPA was associated with an increase of 1.10 scale score in ACT ELA score. In model 2, indicators for having taken advanced coursework in English and social studies were added; this model accounted for 41.3% of the variance in the ACT ELA scores and corresponded to a 9.3% increase in explained variance over model 1. In model 2, having taken advanced coursework in English and social studies was associated with an increase in ACT ELA score of 2.21 and 2.22 scale score points, respectively, after controlling for block 1. The $R^2$ value associated with model 2 is taken as the percentage of variance in ACT ELA scores that is explained by high school grades and coursework taken.

**Table 4.** Blockwise ELA Regression Coefficients

| Predictor | | Model 1 Estimate | Model 2 Estimate | Model 3 Estimate | Model 4 Estimate |
|---|---|---|---|---|---|
| **Intercept** | | 20.12 | 17.98 | 17.21 | 17.67 |
| **English GPA** | | 2.26 | 1.63 | 1.48 | 1.41 |
| **Social Studies GPA** | | 1.10 | 0.88 | 0.77 | 0.73 |
| **Taken Advanced Coursework in English** | | — | 2.21 | 2.15 | 2.22 |
| **Taken Advanced Coursework in Social Studies** | | — | 2.22 | 2.03 | 2.01 |
| **Family Income** | < $36K | — | — | −1.10 | −0.88 |
| | $60K–$100K | — | — | 0.75 | 0.55 |
| | > $100K | — | — | 2.07 | 1.71 |
| | Missing | — | — | 1.14 | 0.95 |
| **Race/Ethnicity** | Asian | — | — | — | 1.63 |
| | Black | — | — | — | −2.72 |
| | Hispanic | — | — | — | −1.24 |
| | Other race/ethnicity | — | — | — | −0.35 |
| | Prefer not to respond/Missing | — | — | — | 0.83 |
| **Gender** | Female | — | — | — | −0.07 |
| | Another Gender/Prefer not to respond/Missing | — | — | — | 2.11 |
| **$R^2$** | | 32.0% | 41.3% | 44.6% | 47.5% |
| **$\Delta R^2$** | | — | 9.3% | 3.3% | 2.9% |

*Note.* English and social studies GPA were standardized. All predictors entered their respective blocks with a significance of <.0001. They were significant at the <.0001 level in the final model.

Model 3 added family income to the previous blocks and accounted for 44.6% of the variance in student ACT ELA scores, which was associated with a 3.3% increase in $R^2$ from model 2. In this model, students with a family income of less than $36,000 had an average ACT ELA score of 1.1 scale score points below students with a family income of $36,000–$60,000. Students with a family income of $60,000–$100,000 had a 0.75 higher ACT ELA score than students with a family income of $36,000–$60,000, after the other variables were accounted for. Students with families whose income was above $100,000 had an average ACT ELA score 2.07 scale score points above that of students with family incomes of $36,000–$60,000, after the other variables were accounted for. Finally, students who did not report a family income had an average ACT ELA score 1.14 scale score points above that of students with family incomes of $36,000–$60,000, after the other variables were accounted for.

Model 4, the full model, explained 47.5% of the variance in ACT ELA scores and was associated with a 2.9% increase in $R^2$ over model 3. This model added race/ethnicity and gender to the model. In this model, students who identified as Asian, Black, Hispanic, and another racial/ethnic category and students who preferred not to respond or did not provide a race/ethnicity had an average ACT ELA score of 1.63 points higher, 2.72 points lower, 1.24 points lower, 0.35 points lower, and 0.83 points, respectively, higher than White students, after the other variables were accounted for. Additionally, female students had an average ACT ELA score 0.07 scale score points lower than male students, and students of another gender, those who preferred not to respond, or those who did not provide their gender had an ACT ELA scale score of 2.11 points higher than male students.

As we can see from the blockwise regression models, English and social studies GPA along with indicators for taking advanced coursework in English and social studies accounted for the largest percentage of variance in ACT ELA scores, 41.3%. The addition of demographic information such as family income, race/ethnicity, and gender only explained an additional 6.2% of the variance in ACT ELA scores.

## Are subgroup differences in ACT STEM and ELA scores reduced after accounting for achievement and academic preparation?

Prior to adjusting for student high school grades and advanced coursework taken, differences in ACT STEM scores between family income levels and racial/ethnic groups were larger than the subgroup differences observed after accounting for student grades, coursework taken, socioeconomic indicators, and demographics (Table 5). For family income, the difference between students with family incomes less than $36,000, $60,000–$100,000, and over $100,000 and students who did not report a family income relative to students whose family income was $36,000–$60,000 was 1.7 points lower, 1.5 points higher, 4.1 points higher, and 1.7 points higher, respectively, on the ACT STEM scores. After adjusting for student grades and coursework taken, these differences were 0.6 points lower, 0.5 points higher, 1.8 points higher, and 1.0 points higher, respectively. For gender, female students had an ACT STEM score 0.8 points lower than male students, and students from another gender, those who preferred not to respond, and those who did not respond had an ACT STEM score of 0.1 points higher than male students prior to adjusting for the full model. After adjusting for the full model, the absolute difference between male and female students increased by 0.6 scale score points, and the difference between male students and students from another gender, those who preferred not to respond, and those who did not respond increased to 0.4 points on the ACT STEM scores. For race/ethnicity, prior to adjusting for student grades and coursework taken, Asian students, Black students, Hispanic students, and students who identified as another race/ethnicity and those who preferred not to respond or who did not respond had a 3.4 point higher, 4.6 point lower, 2.5 point lower, 1.6 point lower, and 1.6 point higher ACT STEM score, respectively, relative to White students. After adjusting for student grades and coursework taking, all subgroup differences were reduced (i.e., 2.0 points higher for Asian students, 2.3 points lower for Black students, 1.3 points lower for Hispanic students, 0.5 points lower for other race/ethnicity, and 0.8 points higher for students who preferred not to respond or missing race/ethnicity).

**Table 5.** Unadjusted and Model Adjusted Mean ACT STEM Scores

| Predictor | | Unadjusted | Δreference group | Full Model | Δreference group |
|---|---|---|---|---|---|
| Family Income | < $36K | 17.5 | −1.7 | 20.4 | −0.6 |
| | $36K–$60K | 19.3 | — | 21.0 | — |
| | $60K–$100K | 20.8 | 1.5 | 21.5 | 0.5 |
| | > $100K | 23.4 | 4.1 | 22.8 | 1.8 |
| | Missing | 21.0 | 1.7 | 22.0 | 1.0 |
| Gender | Female | 20.6 | −0.8 | 20.5 | −1.4 |
| | Male | 21.4 | — | 21.9 | — |
| | Another Gender/Prefer not to respond/Missing | 21.5 | 0.1 | 22.3 | 0.4 |
| Race/Ethnicity | Asian | 24.9 | 3.4 | 23.8 | 2.0 |
| | Black | 17.0 | −4.6 | 19.4 | −2.3 |
| | Hispanic | 19.1 | −2.5 | 20.4 | −1.3 |
| | White | 21.6 | — | 21.8 | — |
| | Other Race/ethnicity | 20.0 | −1.6 | 21.3 | −0.5 |
| | Prefer not to respond/Missing | 23.2 | 1.6 | 22.5 | 0.8 |

*Note.* Reference groups for family income, gender, and race/ethnicity were $36,000–$60,000, male, and White, respectively.

Prior to adjusting for student grades and coursework taken, we can see notable differences in ACT ELA scores between levels of family income, gender groups, and race/ethnicity categories (Table 6). For family income, comparing students from families with household incomes of $36,000–$60,000 and students from families whose income was less than $36,000, $60,000–$100,000, and greater than $100,000 and students who did not provide their family income, there was a difference in ACT ELA scores of −2.0, 1.5, 4.0, and 1.5, respectively. These differences were reduced to −0.9, 0.5, 1.7, and 0.9, respectively, after accounting for the full model. When looking at gender, prior to adjusting for the full model, female students had an ACT ELA score that was 1.1 scale score points higher than male students. Students of another gender, preferred not to respond, or did not respond to the gender question had an ACT ELA score that was 2.3 scale score points above male students. The score difference between male and female students after accounting for student grades and coursework taken was reduced to −0.1 scale score points. However, the difference between male students and students from another gender, who preferred not to respond, or who did not respond remained similar: 2.1. The unadjusted average ACT ELA score differences between White students and Asian students, Black students, Hispanic students, students who identified as another race/ethnicity, and students who preferred not to respond or who did not respond were 3.0, −4.6, −2.3, −1.4, and 1.7 scale score points, respectively. All unadjusted mean differences in comparison to White students were reduced after accounting for student grades and coursework taken (i.e., 1.6 for Asian students, −2.7 for Black students, −1.2 for Hispanic students, −0.4 for other race/ethnicity, and 0.8 for students who preferred not to respond or were missing race/ethnicity). Looking across demographic categories, almost all subgroup differences in ACT ELA scores were reduced after accounting for student grades and coursework taken, except for the comparison of male students to students of another gender, those who preferred not to respond, or those who did not provide their gender.

**Table 6.** Unadjusted and Model Adjusted Mean ACT ELA Scores

| Predictor | | Unadjusted | $\Delta_{\text{reference group}}$ | Full Model | $\Delta_{\text{reference group}}$ |
|---|---|---|---|---|---|
| Family Income | < $36K | 16.3 | −2.0 | 19.1 | −0.9 |
| | $36K–$60K | 18.3 | — | 20.0 | — |
| | $60K–$100K | 19.8 | 1.5 | 20.5 | 0.5 |
| | > $100K | 22.3 | 4.0 | 21.7 | 1.7 |
| | Missing | 19.76 | 1.5 | 20.9 | 0.9 |
| Gender | Female | 20.3 | 1.1 | 19.7 | −0.1 |
| | Male | 19.3 | — | 19.8 | — |
| | Another Gender /Prefer not to respond/Missing | 21.5 | 2.3 | 21.9 | 2.1 |
| Race/Ethnicity | Asian | 23.5 | 3.0 | 22.4 | 1.6 |
| | Black | 15.9 | −4.6 | 18.0 | −2.7 |
| | Hispanic | 18.2 | −2.3 | 19.5 | −1.2 |
| | White | 20.4 | — | 20.7 | — |
| | Other Race/Ethnicity | 19.1 | −1.4 | 20.4 | −0.4 |
| | Prefer not to respond/Missing | 22.1 | 1.7 | 21.6 | 0.8 |

*Note.* Comparison groups for family income, gender, and race/ethnicity were $36,000–$60,000, male, and White, respectively.

## Discussion

This research study explores some of the allegations of socioeconomic, racial/ethnic, and gender biases that are often raised in discussions of the ACT. Notably, the socioeconomic bias argument suggests that students from higher socioeconomic backgrounds are advantaged by their access to superior educational resources, including private tutoring and test preparation courses, which can enhance their performance on the ACT (Kohn, 2000). Another often-raised concern is that of racial or cultural bias, which is centered on inherent biases against certain racial/ethnic groups (FairTest, 2007b). This form of bias argues that cultural references and language nuances may not resonate equally across all student populations. Finally, the concern of gender bias, particularly in the math and science sections of the ACT, has raised questions about the test's ability to fairly predict academic outcomes for all students (FairTest, 2007a).

To preemptively address these concerns, ACT has implemented a number of measures aimed at minimizing potential biases and ensuring an equitable assessment of all students. These efforts include engaging a diverse group of individuals in the test content development process, conducting external fairness reviews, and employing statistical analysis to identify and address differential item functioning. Additionally, socioeconomic disparities are addressed through social programs such as fee waivers and free test preparation resources, which strive to level the playing field for students from all backgrounds.

The study by McNeish et al. (2015) offers an important perspective on the factors influencing ACT Composite scores and highlights the role of academic preparation while noting the relatively minor role of socioeconomic and demographic factors after adjusting for student and school characteristics. This study also highlights the importance of holistic evaluation of student performance, including the consideration of academic factors. The current research extends this understanding by exploring ACT STEM and ACT ELA scores and aims to illustrate that disparities observed between student subgroups could diminish significantly once we account for student high school grades and advanced coursework taken, thereby furthering the discussion on standardized testing fairness and equity.

In the case of both ACT STEM and ACT ELA scores, students' high school grades and advanced coursework taken explained the largest proportion of variance in these scores. Students' socioeconomic status and demographics accounted for significantly less of the explained variance in these scores once student grades and coursework were accounted for. In fact, for ACT STEM and ACT ELA scores, socioeconomic status and demographic characteristics only accounted for an increase of 8.4 and 6.2 percent of the variance in each score respectively. In contrast, students' high school grades and advanced coursework taken accounted for 41.8 and 41.3 percent of the variance in ACT STEM and ACT ELA scores, respectively.

Additionally, accounting for student high school grades and advanced coursework taken resulted in substantial decreases in subgroup differences for both ACT STEM and ACT ELA scores by family income and race/ethnicity. While there was a decrease in subgroup differences between male and female students in the ACT ELA scores, this was not observed for ACT STEM scores. The median standard error of measurement for the ACT STEM and ACT ELA tests were 1.26 and 1.43, respectively. The standard error of measurement provides an estimate of the range within which an individual's true score likely lies. It helps to quantify the uncertainty of an individual test score and offers an interpretation of test scores by providing a confidence interval around the obtained score.

A student's given examination score is only an estimate of that examinee's true scale score. The true score can be interpreted as the average score obtained over countless repeated administrations of the test under identical conditions. When viewed in terms of groups of examinees, if one standard error of measurement was added and subtracted to the reported score for each examinee, the resulting intervals would contain the true scores for approximately 68% of the examinees. Given the standard error of measurement for the ACT STEM and ACT ELA tests and the adjusted subgroup differences observed in this study, most of the subgroup differences were within the standard error of measurement, while some differences between subgroups were larger than would be accounted for by the standard error of measurement.

The existence of subgroup differences in and of themselves does not indicate that there is bias in the test, as argued by some external individuals. As can be observed by the methodology employed in this study, accounting for student characteristics can reduce observed subgroup differences. By this logic, the inclusion of additional student, school, and environmental factors, such as the number of advanced courses taken or the grades in advanced coursework, could further reduce subgroup differences. This study highlights that accounting for student grades and coursework taken explains much of the variance in ACT STEM and ACT ELA scores and that the additional variables, while explanatory in nature, have less of an impact on ACT STEM and ACT ELA scores.

# References


Camara, W., Kimmel, E., Scheuneman, J., & Sawtell, E. A. (2003). *Whose grades are inflated?* (Research Report No. 2003-4). College Board.

Clarke, P. (2008). When can group level clustering be ignored? Multilevel models versus single-level models with sparse data. *Journal of Epidemiology and Community Health, 62*(8), 752–758.

Clarke, P., & Wheaton, B. (2007). Addressing data sparseness in contextual population research: Using cluster analysis to create synthetic neighborhoods. *Sociological Methods & Research, 35*(3), 311–351.

FairTest. (2007a, August 20). *Gender bias in college admissions tests* [press release]. https://fairtest.org/gender-bias-college-admissions-tests/

FairTest. (2007b, August 20). *The ACT: Biased, inaccurate, and misused.* https://fairtest.org/act-biased-inaccurate-and-misused/#:~:text=One%20study%20conducted%20at%20a,approximately%2028%25%20of%20the%20differences.

Kohn, A. (2000). Fighting the tests: A practical guide to rescuing our schools. *Cultural Logic: A Journal of Marxist Theory & Practice*, *7*(2000).

Kuncel, N. R., Credé, M., & Thomas, L. L. (2005). The validity of self-reported grade point averages, class ranks, and test scores: A meta-analysis and review of the literature. *Review of Educational Research, 75*(1), 63–82.

McNeish, D. M. (2014). Modeling sparsely clustered data: Design-based, model-based, and single-level methods. *Psychological Methods*, *19*(4), 552.

McNeish, D. M., Radunzel, J., & Sanchez, E. (2015). *A multidimensional perspective of college readiness: Relating student and school characteristics to performance on the ACT.* https://www.act.org/content/dam/act/unsecured/documents/ACT_RR2015-6.pdf

Milner, H. R., IV. (2013). Rethinking achievement gap talk in urban education. *Urban Education*, *48*(1), 3–8.

Sanchez, E. I., & Buddin, R. (2015). *How accurate are self-reported high school courses, course grades, and grade point average?* ACT.

Shaw, E. J., & Mattern, K. D. (2009). Examining the accuracy of self-reported high school grade point average. (Research Report No. 2009-5). College Board.

# Appendix

**Table A1.** Percentages of the STEM Population and the STEM Sample Used in the Study

| Characteristic | | Population | | Sample | |
|---|---|---|---|---|---|
| | | N | Percent | N | Percent |
| Race/Ethnicity | Asian | 44,414 | 5% | 16,907 | 5% |
| | Black | 93,267 | 11% | 35,349 | 11% |
| | Hispanic | 116,574 | 13% | 44,181 | 13% |
| | White | 547,767 | 62% | 207,642 | 62% |
| | Other | 50,980 | 6% | 19,375 | 6% |
| | Prefer not to respond/Missing | 24,915 | 3% | 9,546 | 3% |
| Gender | Female | 472,268 | 54% | 179,200 | 54% |
| | Male | 394,790 | 45% | 149,587 | 45% |
| | Another Gender/Prefer not to respond/Missing | 10,859 | 1% | 4,213 | 1% |
| Family Income | < $36K | 138,008 | 16% | 52,221 | 16% |
| | $36K–$60K | 108,370 | 12% | 41,120 | 12% |
| | $60K–$100K | 152,228 | 17% | 57,706 | 17% |
| | > $100K | 312,896 | 36% | 118,374 | 36% |
| | Missing | 166,415 | 19% | 63,579 | 19% |
| Race/Ethnicity*Gender*Family Income | Asian*Female*<$36K | 3,517 | 0% | 1,327 | 0% |
| | Asian*Female*$36K–$60K | 2,791 | 0% | 1,110 | 0% |
| | Asian*Female*$60K–$100K | 3,629 | 0% | 1,366 | 0% |
| | Asian*Female*>$100K | 9,306 | 1% | 3,540 | 1% |
| | Asian*Male*<$36K | 2,592 | 0% | 1,025 | 0% |
| | Asian*Male*$36K–$60K | 2,460 | 0% | 955 | 0% |
| | Asian*Male*$60K–$100K | 3,113 | 0% | 1,144 | 0% |
| | Asian*Male*>$100K | 8,229 | 1% | 3,106 | 1% |
| | Black*Female*<$36K | 20,811 | 2% | 7,951 | 2% |
| | Black*Female*$36K–$60K | 9,649 | 1% | 3,612 | 1% |
| | Black*Female*$60K–$100K | 6,829 | 1% | 2,583 | 1% |
| | Black*Female*>$100K | 6,191 | 1% | 2,361 | 1% |
| | Black*Male*<$36K | 12,348 | 1% | 4,626 | 1% |
| | Black*Male*$36K–$60K | 7,414 | 1% | 2,832 | 1% |
| | Black*Male*$60K–$100K | 6,099 | 1% | 2,214 | 1% |
| | Black*Male*>$100K | 5,900 | 1% | 2,291 | 1% |
| | Hispanic*Female*<$36K | 19,586 | 2% | 7,322 | 2% |
| | Hispanic*Female*$36K–$60K | 11,047 | 1% | 4,261 | 1% |
| | Hispanic*Female*$60K–$100K | 9,422 | 1% | 3,504 | 1% |
| | Hispanic*Female*>$100K | 11,829 | 1% | 4,547 | 1% |
| | Hispanic*Male*<$36K | 11,930 | 1% | 4,501 | 1% |
| | Hispanic*Male*$36K–$60K | 8,628 | 1% | 3,302 | 1% |
| | Hispanic*Male*$60K–$100K | 8,114 | 1% | 3,078 | 1% |
| | Hispanic*Male*>$100K | 11,091 | 1% | 4,174 | 1% |
| | White*Female*<$36K | 32,179 | 4% | 12,131 | 4% |
| | White*Female*$36K–$60K | 31,420 | 4% | 11,854 | 4% |
| | White*Female*$60K–$100K | 54,145 | 6% | 20,599 | 6% |

| Characteristic | | Population | | Sample | |
|---|---|---|---|---|---|
| | | N | Percent | N | Percent |
| | White*Female*>$100K | 120,246 | 14% | 45,451 | 14% |
| | White*Male*<$36K | 20,473 | 2% | 7,825 | 2% |
| | White*Male*$36K–$60K | 24,192 | 3% | 9,086 | 3% |
| | White*Male*$60K–$100K | 47,270 | 5% | 17,990 | 5% |
| | White*Male*>$100K | 114,382 | 13% | 43,156 | 13% |
| | Missing*Missing*Missing | 231,085 | 26% | 88,176 | 26% |
| **Race/Ethnicity*Gender** | Asian*Female | 24,288 | 3% | 9,250 | 3% |
| | Asian*Male | 19,862 | 2% | 7,565 | 2% |
| | Black*Female | 53,041 | 6% | 20,132 | 6% |
| | Black*Male | 39,785 | 5% | 15,035 | 5% |
| | Hispanic*Female | 65,536 | 7% | 24,799 | 7% |
| | Hispanic*Male | 49,880 | 6% | 18,938 | 6% |
| | White*Female | 291,142 | 33% | 110,388 | 33% |
| | White*Male | 251,027 | 29% | 95,099 | 29% |
| | Missing*Missing | 83,356 | 9% | 31,794 | 10% |
| **Race/Ethnicity*Family Income** | Asian*<$36K | 6,140 | 1% | 2,364 | 1% |
| | Asian*$36K–$60K | 5,295 | 1% | 2,077 | 1% |
| | Asian*$60K–$100K | 6,794 | 1% | 2,529 | 1% |
| | Asian*>$100K | 17,604 | 2% | 6,675 | 2% |
| | Black*<$36K | 33,310 | 4% | 12,639 | 4% |
| | Black*$36K–$60K | 17,141 | 2% | 6,475 | 2% |
| | Black*$60K–$100K | 12,995 | 1% | 4,835 | 1% |
| | Black*>$100K | 12,138 | 1% | 4,663 | 1% |
| | Hispanic*<$36K | 31,856 | 4% | 11,944 | 4% |
| | Hispanic*$36K–$60K | 19,865 | 2% | 7,633 | 2% |
| | Hispanic*$60K–$100K | 17,708 | 2% | 6,644 | 2% |
| | Hispanic*>$100K | 23,081 | 3% | 8,783 | 3% |
| | White*<$36K | 53,687 | 6% | 20,369 | 6% |
| | White*$36K–$60K | 56,460 | 6% | 21,281 | 6% |
| | White*$60K–$100K | 102,403 | 12% | 38,956 | 12% |
| | White*>$100K | 235,995 | 27% | 89,113 | 27% |
| | Missing*Missing | 225,445 | 26% | 86,020 | 26% |
| **Gender*Family Income** | Female*<$36K | 83,499 | 10% | 31,536 | 9% |
| | Female*$36K–$60K | 59,996 | 7% | 22,791 | 7% |
| | Female*$60K–$100K | 80,127 | 9% | 30,423 | 9% |
| | Female*>$100K | 159,025 | 18% | 60,224 | 18% |
| | Male*<$36K | 52,446 | 6% | 19,872 | 6% |
| | Male*$36K–$60K | 46,828 | 5% | 17,732 | 5% |
| | Male*$60K–$100K | 70,382 | 8% | 26,627 | 8% |
| | Male*>$100K | 151,351 | 17% | 57,197 | 17% |
| | Missing*Missing | 174,263 | 20% | 66,598 | 20% |
| **TOTAL** | | 877,917 | 100% | 333,000 | 100% |

## Notes

[1] $36,000 to $60,000 was selected as the reference group as it represented a lower middle-income category.

## ABOUT ACT

ACT is transforming college and career readiness pathways so that everyone can discover and fulfill their potential. Grounded in more than 65 years of research, ACT's learning resources, assessments, research, and work-ready credentials are trusted by students, job seekers, educators, schools, government agencies, and employers in the U.S. and around the world to help people achieve their education and career goals at every stage of life. Visit us at www.act.org.