Annual Changes in High School Average ACT Composite Scores: Celebration, Concern, or Something Else?

Matt T. Harmston, MA, Wayne J. Camara, PhD, and Christine K. Phillips

Schools, districts, and states frequently use ACT® scores as indicators of school quality and progress over time. Small increases in school-level ACT outcomes are often met with celebration, or when they decrease, with concern. To facilitate assignment of meaning to such change, ACT has quantified school-level changes in outcomes by enabling comparison to one-year school-level change distributions. Crouse and Harmston (2018) quantified changes in percentage meeting three or four ACT College and Career Readiness Benchmarks. Using similar methodology, Harmston and Crouse (2016) enabled education leaders to identify the percentile rank of their school's annual change in average Composite scores compared to a nationwide set of similar sized schools. Though earlier papers provided a framework with which to better understand change, this paper moves a step further by providing context and an online application (see sidebar) to aid the interpretation of annual changes in school-level average ACT Composite scores relative to a five-year national comparison group.

When educators attempt to understand and interpret changes in aggregate ACT scores, they often have a number of questions:

- Are ACT scores representative of all students?
- What are some considerations for institutional- and state-level ACT performance comparisons?
- Do annual changes in a school's average test scores indicate a substantive difference in student performance?
- How do local changes in annual average ACT test scores compare to schools nationwide?
- What do changes in average test scores tell us about school quality and progress?

Average Score Change Percentile Rank Tool

We developed an interactive tool that automatically computes a school's change in average ACT Composite score and associated percentile rank compared to a national comparison group of schools testing similar numbers of graduates.

The tool allows users to quickly identify average score change and percentile rank, by selecting both current and prior year average Composite scores, as well as number of ACT-tested graduates for the current year (found in the executive summary of the school's graduating class ACT Profile Report). After selections are made, the tool automatically populates the school's score change, percentile rank, and additional interpretive text.

See tool

Are ACT test scores representative of all students?

In 2017-18, the ACT was administered to nearly all public-school juniors in 19 states, a model referred to as "census-testing."[1] Similarly, individual schools and districts in other states may also choose to administer the ACT to all students in a class (e.g., 11th graders). Aggregated ACT scores and other data from schools, districts, and states participating in such census-testing models are representative of all graduates in respective high schools because nearly 100% of students participate in the testing. In the remaining schools, districts, and states, students and parents "self-select" whether a student will complete an admissions test (ACT or otherwise). In these situations, participation rates for 2019 high school graduates will vary greatly.

As indicated above, aggregate (averaged) scores in schools, districts, and states using a census-testing model can be accurately characterized as being representative of all students in the population from which they came. However, in jurisdictions where students self-select to take the ACT (or admissions tests in general), aggregate outcomes often do not represent the student-body as a whole. Rather, aggregate outcomes for self-selected cohorts differ in meaningful ways from groups who do not participate in admissions testing. Self-selected ACT-testing students are frequently college-bound, and as such, may differ from non-college-bound cohorts (and the overall student population) in terms of higher grades, more rigorous coursework, higher socio-economic backgrounds, and other factors. In association with these factors, we typically observe that college-bound students have higher test scores. These factors should limit interpretations about aggregate performance (e.g., school quality, growth) when participation rates fall substantially below 100%.

Implementation of ACT census-testing facilitates interpretability, given that test scores (and their changes over time) reflect entire populations (i.e., all public high school 11th graders and their respective graduating classes). When participation rates dip below 90-100%, changes to the student population impact the manner in which outcomes are generalizable to the entire student body. However, any school, district, or state enrolling ACT test-takers can still find value in examining aggregate trend data across multiple years so long as they are aware of the appropriate limitations and constraints. Questions to be considered when examining admission score trends and year-to-year change include:

- Has the percentage of students taking the ACT changed significantly compared to prior years?
- Have graduating classes (and the locale from which they came)[2] changed significantly in terms of size, academic preparation, and other factors which could influence performance?
- Could changes in curriculum or policy have affected performance?
- How does graduating class performance compare to that of similar schools/districts/states?[3]
- How are students performing on the ACT relative to performance on other tests (e.g., state high school assessments)?

To reiterate, interpretability of school-level ACT outcomes is enhanced when institutions adopt all-inclusive participation in ACT testing. Still, even if schools are not testing all students, score change data presented in this brief will be useful for any school that remains fairly stable in terms of participation rate and student background characteristics. Regardless of testing model, monitoring change in ACT outcomes helps guide educators' efforts to facilitate students' pursuit of excellence in education.

What are some considerations for institutional- and state-level ACT performance comparisons?

Comparisons between schools, districts, and states should be conducted with great care even when ACT census-testing has been implemented. Non-academic differences between institutions in terms of size, location, student backgrounds, institutional resources,

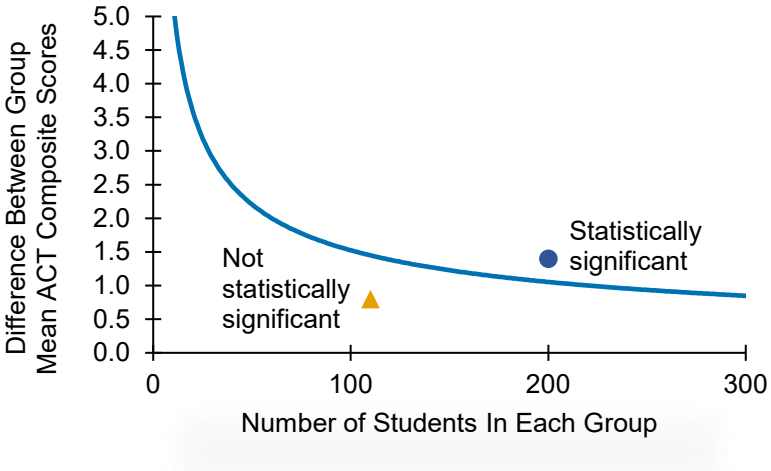
curricular changes, etc. may impact student performance. Another type of comparison, and the type of primary interest for this paper, is considered within the institution. Within-school comparisons focus on changes in aggregate performance for students attending the same school. For instance, schools that have been administering the ACT to all students for three or more years typically have consistent populations over time. As a result, schools can more readily discern the degree to which their average scores have changed across time due to academic factors. This point also applies to districts and states that test all students for several consecutive years.

The degree to which scores change upon retesting is often an important consideration for individual students, such that counselors wish to determine where students exhibited the most gain. In such an instance, the standard error of measurement (SEM) should be considered when comparing differences between test scores for an individual student or between students.[4] A similar type of variation exists across scores when they are averaged for a school, district, or state. Taking this variation into account is particularly important when making inferences about the statistical or practical significance of year-to-year changes in average scores. These and other cautions apply to interpretation of aggregate scores and score changes for any assessment program, not just the ACT.

Do annual changes in a school's average test scores indicate a substantive difference in student performance?

When considering the meaning of change, attention often turns to the question of statistical significance. Statistical significance depends on a variety of factors, among them the number of students tested and the magnitude of that change. A one-point difference may be statistically significant in schools testing a large number of students, whereas that same difference may not be statistically significant in a school testing fewer students. Allen (2016) provided an illustration to assist with determining significance of differences between average ACT Composite scores for two groups as a function of size (see Figure 1). The groups could be schools, districts, or other manner of grouping students.

Figure 1. Estimated Statistical Significance of ACT Composite Score Mean Differences



How to use this graph:

First, determine the number of students testing in each group. Then, average those numbers. Locate that point on the horizontal axis.

Next, locate the difference in average ACT Composite scores on the vertical axis.

Finally, locate the point where both values intersect on the graph. Points above the curve (e.g., blue circle) are **likely** to be statistically significant. Points below (e.g., orange triangle) are **unlikely** to be statistically significant (Allen, 2016).

Does an absence of statistical significance indicate an absence of substantive change?

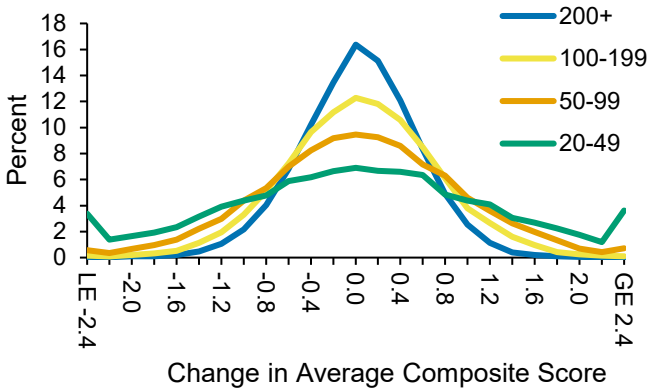
While statistical significance has a distinct allure, sometimes meaningful change isn't statistically significant. When viewing year-to-year school averages, it is tempting to ascribe undue importance to small score changes. When considering such change, it is important to remember that variability in aggregate test scores is impacted by sample size. For instance, a school with 15 graduates taking the ACT is likely to experience greater year-to-year change in average ACT scores than a school testing 85 graduates. With this issue in mind, examining score trends (e.g. 5 years, 10 years) can provide valuable information in addition to year-to-year change. Variation in average scores will still be associated with number of graduates tested, but changes can be more readily linked to factors endemic to the local educational system.

How do local changes in annual average ACT test scores compare to schools nationwide?

Given that there are conditions that bring about greater variability in a school's aggregate performance, institutions need a way to quantify how their yearly change compares to others similar in size. In replicating methodology by Harmston and Crouse (2016), we also expanded upon their paper by analyzing data on a reference group composed of year-to-year changes in high school average ACT Composite scores for 2013 through 2018 graduating classes. Using five years of aggregate difference scores increased the stability of findings, thereby providing a consistent reference point for schools to better understand changes in average scores for any recent pair of adjacent years.

As noted above, schools with fewer test-takers will, on average, see greater year-to-year score variability than schools with more test-takers. To illustrate differences in average score variability due to sample size, Figure 2 provides distributions of year-to-year score changes for high schools broken out by number of ACT-tested graduates. Differences in variability are illustrated by smaller schools having a flatter, wider profile, compared to more centralized profiles of progressively larger schools.

Figure 2. Distributions of Change in Average Composite Scores



Technical Note and Methods

We generated cumulative percent distributions of average ACT Composite change scores for high schools falling into specific size ranges based on the number of students taking the ACT, bracketed by counts of: 20-49 (12,727 difference scores), 50 to 99 (12,953 difference scores), 100 to 199 (11,828 difference scores), and greater than 200 (11,660 difference scores). A large change in the number of students testing with the ACT between adjacent years is often an indication of change within a school's student population. To control for extreme variation in number tested, difference scores were ranked according to their associated percent change in number of students tested. Final analyses **excluded** change scores from instances where the percent-change in n-count fell in the top and bottom quarters of respective distributions. This control mitigated some of the impact of changes in contract testing participation over time.

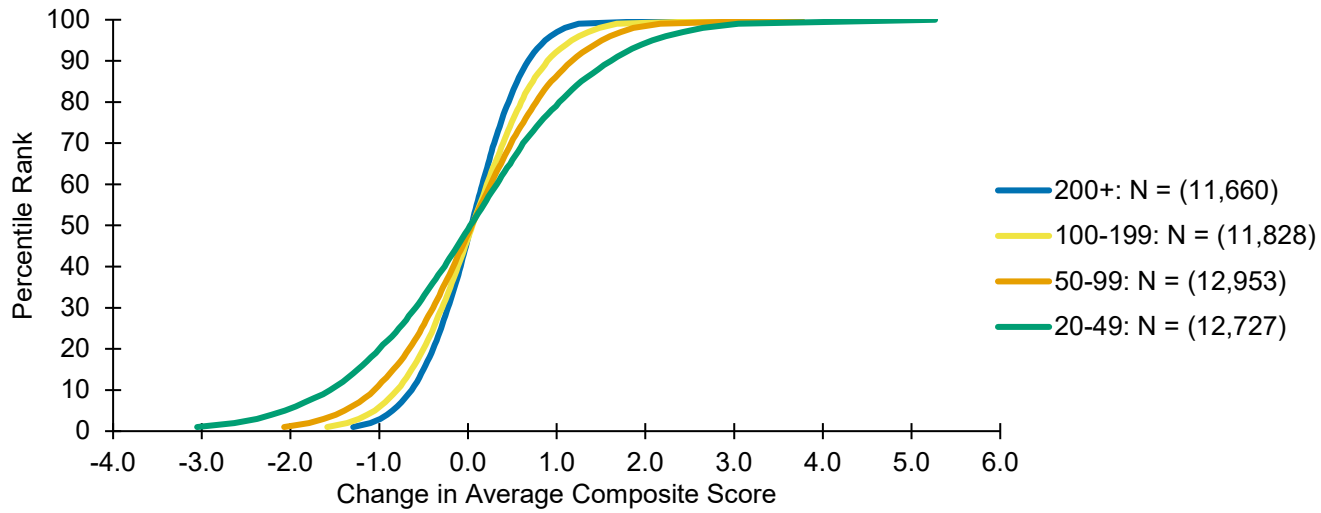
For a coarse estimate of percentile ranks associated with average score change, see Table 1. The table is based upon the same source data used for Figure 2. First, locate the rounded score change value in the first column of Table 1 closest to your school's average Composite score change. Then, scan across that row until you find the column corresponding to numbers of ACT-tested students in the most recent graduating class. The corresponding cell value represents a rough estimate of the percentage of schools demonstrating score change levels the same as or lower than your school. Schools with less than 20 students should use the "20-49" column.

If you seek a more specific percentile rank, detailed cumulative percent distributions are plotted in the curves found in Figure 3. First, find the curve representing the group into which your school's present-year number of ACT-tested graduates falls. Then, find the rough location of this year's change in average ACT Composite score on the horizontal axis, drawing a vertical line up to your sample-size curve. From the point where your curve and the vertical line intersect, draw a horizontal line to the vertical axis. This point of intersection represents the percentage of difference scores falling at or below your change in average composite score (e.g., your percentile rank). As with Table 1, schools with less than 20 ACT-tested graduates should compare to the "20-49" curve.

Table 1. Percentile Ranks of Select One-Year School-Level Average ACT Composite Score Changes by Number of Graduates Tested

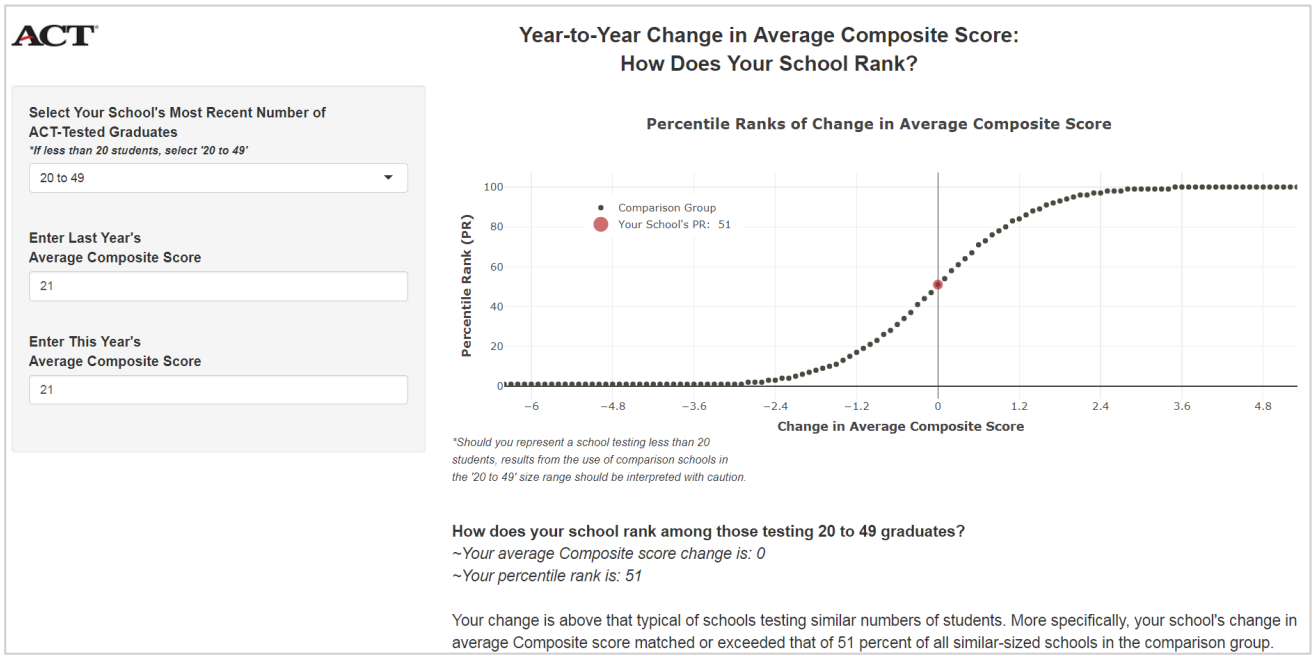
	School Size			
Change	**20-49**	**50-99**	**100-199**	**200+**
LE -2.4	3	1	0	0
-2.2	5	1	0	0
-2.0	6	2	0	0
-1.8	8	3	1	0
-1.6	11	4	1	0
-1.4	14	6	2	1
-1.2	18	9	4	2
-1.0	22	13	8	4
-0.8	27	19	13	8
-0.6	33	26	20	15
-0.4	39	34	30	25
-0.2	46	43	41	39
0.0	**53**	**53**	**53**	**55**
0.2	59	62	65	70
0.4	66	71	76	82
0.6	72	78	84	91
0.8	77	84	90	96
1.0	81	89	94	98
1.2	85	92	97	99
1.4	89	95	98	100
1.6	91	97	99	100
1.8	93	98	100	100
2.0	95	99	100	100
2.2	96	99	100	100
GE 2.4	100	100	100	100

Figure 3. One-Year Change in School-Level Average ACT Composite Score by Number of Graduates Tested



With outcomes presented in the graphical format of Figure 3, steeply-sloped curves and extreme increases for small schools introduce challenges for determining a school's exact percentile rank. It is with these challenges in mind that we built the tool highlighted at the beginning of this paper. Our primary goal was to enable users to easily identify schools' specific percentile ranks, eliminating the need for chart- and table-based estimation (see Figure 4 below).

Figure 4. Screen-Capture of Interactive Tool for Determining Percentile Rank of Yearly Change in School-Level Average ACT Composite Scores



The intent for presenting multiple options to find a school's rank is to enable users to more readily interpret and understand a school's changes in average ACT Composite score.

What do changes in average test scores tell us about school quality and progress?

We have discussed factors related to change in aggregate test scores on the ACT and provided illustrations to assist with deriving meaning in magnitude of change relative to other schools. In doing so, this information should reduce misinterpretation of annual average score variation. However, magnitude of change begs the question of what such a shift can ultimately tell us about school quality and progress.

ACT scores provide useful, reliable metrics for evaluating students' college and career preparation and readiness. However, shifts in aggregate performance need to be interpreted in context: Not all change is due to academic factors, thus may not reflect changes in quality or progress. As described above, a multitude of factors lead to score change that may or may not be academic in nature. Adding to the complexity is random variation. Only after understanding the impact of non-academic factors can schools more readily discern how much of their score change is due to academics. This understanding can be improved upon by supplementing year-to-year change inquiries with multi-year trend data. Consistently testing a high percentage of a school's students further strengthens conclusions regarding quality and progress.

Regardless of testing model, it is important to remember that year-to-year change in test scores—whether using the ACT or another state/national test—represents but one necessary piece of a larger evidentiary puzzle depicting a school's quality and progress. This piece, and the puzzle as a whole, are best reviewed in context.

Notes

1. Thirteen states administered the ACT to all 11th graders (AL, AR, KY, LA, MS, MT, NE, NV, NC, ND, UT, WI, WY); 6 states offered district choice (AZ, HI, OH, OK, SC, TN) under which a large majority of students completed the ACT.
2. This issue is particularly relevant to districts with two or more high schools and communities that share common schools when changing boundaries may alter the composition of the student body.
3. Many factors should be considered, including school size, participation rates, socioeconomic factors (e.g., parental education, eligibility for free or reduced school lunch), demographic backgrounds, etc.
4. SEM: The degree to which observed scores vary around a theoretical "true score."

References

Allen, J. (2016). *Comparing average ACT Composite scores for two groups*. Iowa City, IA: ACT. Retrieved from https://www.act.org/content/dam/act/unsecured/documents/6575_Tech-Brief-Comparing-Ave-Composite-Scores.pdf

Crouse, J., & Harmston, M. (2018). *Interpreting one-year changes in college readiness*. Iowa City, IA: ACT. Retrieved from https://www.act.org/content/dam/act/unsecured/documents/pdfs/R1691-interpreting-cr-changes-2018-10.pdf

Harmston, M., & Crouse, J. (2016). *Interpreting one-year average score changes*. Iowa City, IA: ACT. Retrieved from https://www.act.org/content/dam/act/unsecured/documents/ACT-One-Year%20Average-Score-Changes.pdf

Matt T. Harmston, MA

Matt Harmston is a senior research scientist in Research Technology, Data Science, and Analytics specializing in secondary research and report development.

Wayne J. Camara, PhD

Wayne Camara is ACT's Horace Mann Chair. He has served as SVP of Research at ACT, Vice President of R&D at College Board and Associate Executive Director of Science at the American Psychological Association. His research spans issues of validity, fairness, and utility of assessments in admissions, accountability, and employment.

Christine K. Phillips

Christine Phillips is a program manager in Research Technology, Data Science, and Analytics, focusing on custom research projects and ACT contract reporting.