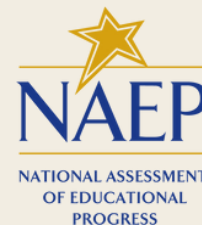


INTERPRETING NAEP RESULTS

Five Reminders from the National Center for Education Statistics (NCES)



Check for statistical significance.

NCES uses statistical significance tests with a 95 percent confidence level to evaluate differences in NAEP results. Others may use different confidence levels (e.g., 90 percent) or methods (e.g., Bayesian). Be sure to use statistical tests when comparing results over time, across jurisdictions, and among student groups.

* Significantly different ($p < .05$) from 2024.

Mathematics 2024

Grade 4

Difference from Michigan		
Colorado	239	↑
Connecticut	239	◆
South Carolina	238	◆
Michigan	235	—

↑ Significantly higher

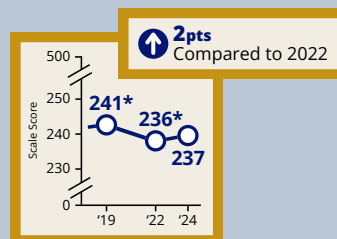
◆ Not significantly different

Rely on NAEP reports and tools to compare results across states and districts.

NCES clearly indicates whether jurisdictions' results are significantly higher, lower, or not statistically different from others. Do not simply compare average score estimates; use the indicators of statistical significance provided by NCES.

Remember that NCES calculates differences based on unrounded numbers but reports findings as rounded numbers.

Example: What looks like a 1-point difference in the grade 4 mathematics results in 2022 and 2024 is actually a 2-point difference based on the unrounded results from each year.



Grade 4

Mathematics 2024

Mathematics 2022

Grade 4

13 states/jurisdictions had standards that mapped to *NAEP Proficient*

37 states/jurisdictions had standards that mapped to *NAEP Basic*

1 state/jurisdiction had standards that mapped to below *NAEP Basic*

Emphasize that *NAEP Proficient* does not always align with states' grade-level performance standards.

The NAEP standard for proficiency represents competency over challenging subject matter, a standard that exceeds most states' standards for proficient or grade-level achievement.

Use caution when examining relationships between student experiences and achievement.

Although NAEP results compare student performance based on demographic characteristics and educational experiences, the comparisons cannot be used to establish cause-and-effect relationships. Many factors may influence academic achievement.

Correlation
≠
Causation