

# **The National Assessment of Educational Progress (NAEP):**

## **A Beat Reporter's Secret Weapon**

---

Presented by the Center on Education Policy  
and the Education Writers Association

# What is NAEP?

- Created in mid-1960s to be a common, national barometer of US student achievement
- Overseen and funded by the National Center on Education Statistics (NCES) but governed by an independent board
- Assesses a representative sample of students in grades 4, 8 and 12
- Assesses reading and math every two years, other subjects less frequently

# How are NAEP results reported?

- By average scale scores *and*
- By percentages of students reaching 3 levels of achievement
  - Basic, proficient, advanced
- By student subgroups
  - Major racial/ethnic groups
  - Gender
  - Income (school lunch eligibility)
  - Other variables

# Long-term trend NAEP

- Special assessment that measures long-term trends in student achievement
  - Differs in content from “main” NAEP
  - Reports trends since 1970s
  - Covers reading and math only
  - Is administered every 4 years based on age (not grade) to students ages 9, 13 and 17

# What makes NAEP unique as a test?

- It tests only a *sample* of students and schools
- It provides no individual student scores or school results (low consequences)
- Results are reported only for the nation, states and large urban districts (TUDA)
- It's fast! Student testing time of less than an hour

# Key details to remember about NAEP

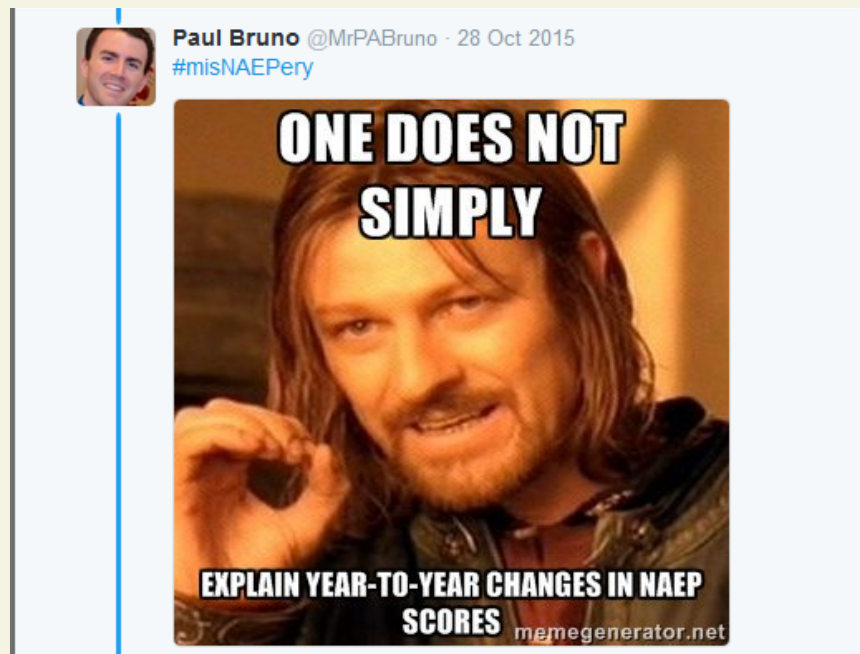
- NAEP content is NOT aligned to any state's curriculum, so it may not always test what kids are being taught in school.
- NAEP's definition of "proficient" is a high bar (much higher than grade-level performance and most state proficiency levels)

# Key details to remember about NAEP

- Not all differences are statistically significant.
- Short-term *changes* in results are not as revealing as longer-term *trends*. A drop or gain between one testing cycle is not a trend.
- Demographic changes in the student population can affect average scores over the long term
  - Students from historically underperforming groups make up a larger share of the population than they did 40 years ago

# MisNAEPery is a thing...

- Misreporting of NAEP happens so often it has become a meme (albeit a nerdy one).





# Causation vs. correlation

- Attributing drops or gains in NAEP to any one “cause” is a common problem.
  - *Causes* may be influenced by several factors, including variables that cannot be measured.
- *Correlations* between the data and other variables are important to consider, but citing a direct causal relationship is an inappropriate use of the data.

# The 2015 NAEP data: A case study for responsible correlating

*“Widespread adoption of the Common Core was responsible for the 2015 decline in NAEP scores in reading and math!”*

Why is this statement misNAEPery?

# MisNAEPery exposed!

The statement is inaccurate because:

- Education systems are complex. It is impossible to attribute any changes in NAEP to one discrete variable, in this case the Common Core.
- The decline in 2015 NAEP scores was likely influenced by a range of variables.
- It is fair, however, to consider which variables might be the most significant.

# Considering significant correlations

- Because the adoption of the Common Core led to significant changes in school and classroom practices, it is reasonable to assume some of those changes may have impacted NAEP results.
- Other credible research confirms the impact of these changes. CEP and others have reported on “implementation chaos” within schools and districts as a result of the new standards.

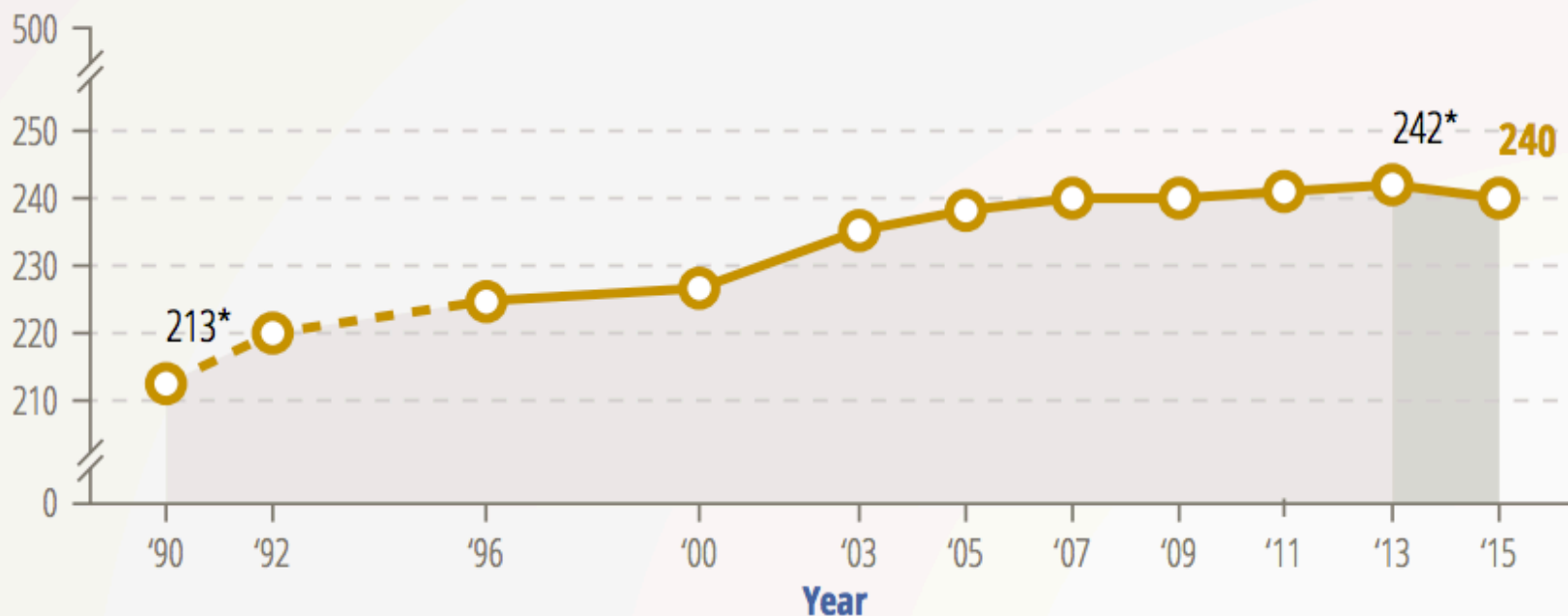
## Other common types of misNAEPery

- Cherry picking — focusing only on subjects & grades that support your point
- Comparing scores between grades 4, 8 and 12 — the scoring scales for different grades are not comparable
- Using NAEP “proficiency” as a proxy for grade-level or acceptable performance

# Telling the real story: 2015 NAEP scores & trends

## Mathematics, grade 4

Scale Score

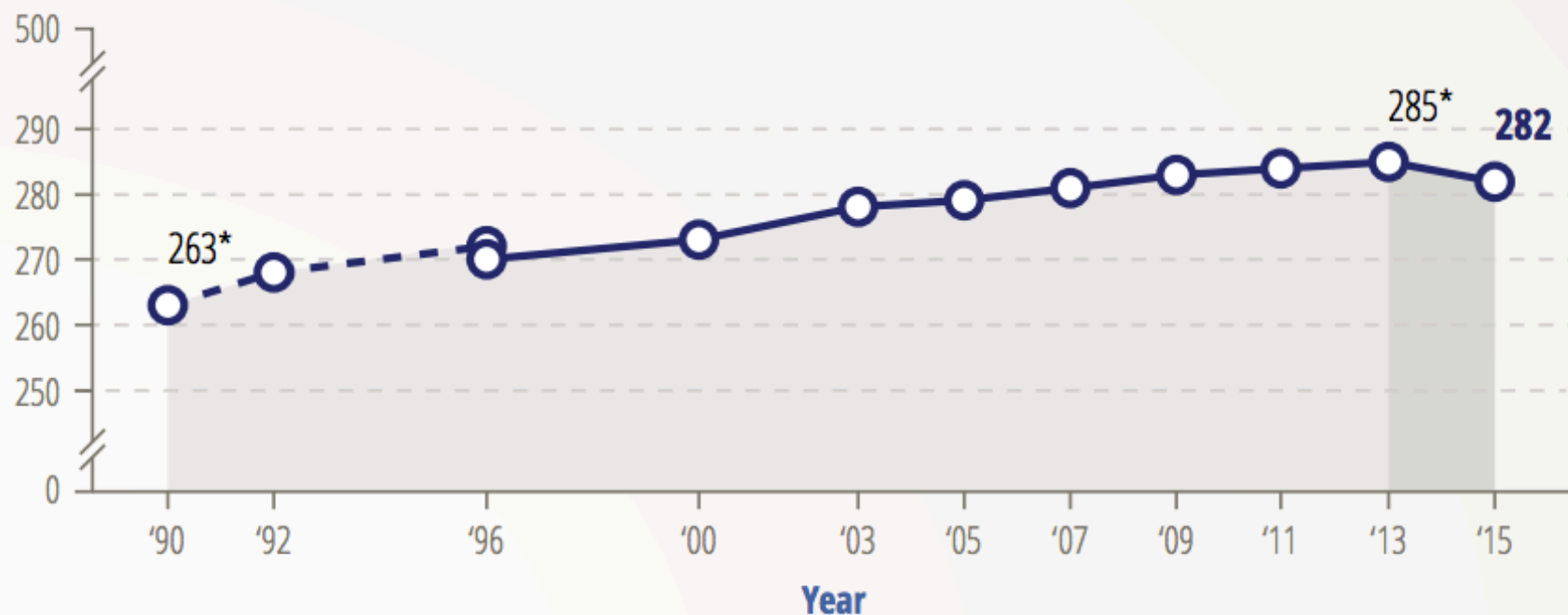


The dashed line indicates years in which NAEP was administered and accommodations were not available for students taking the assessment.

\* Significantly different ( $p < .05$ ) from 2015

# Mathematics, grade 8

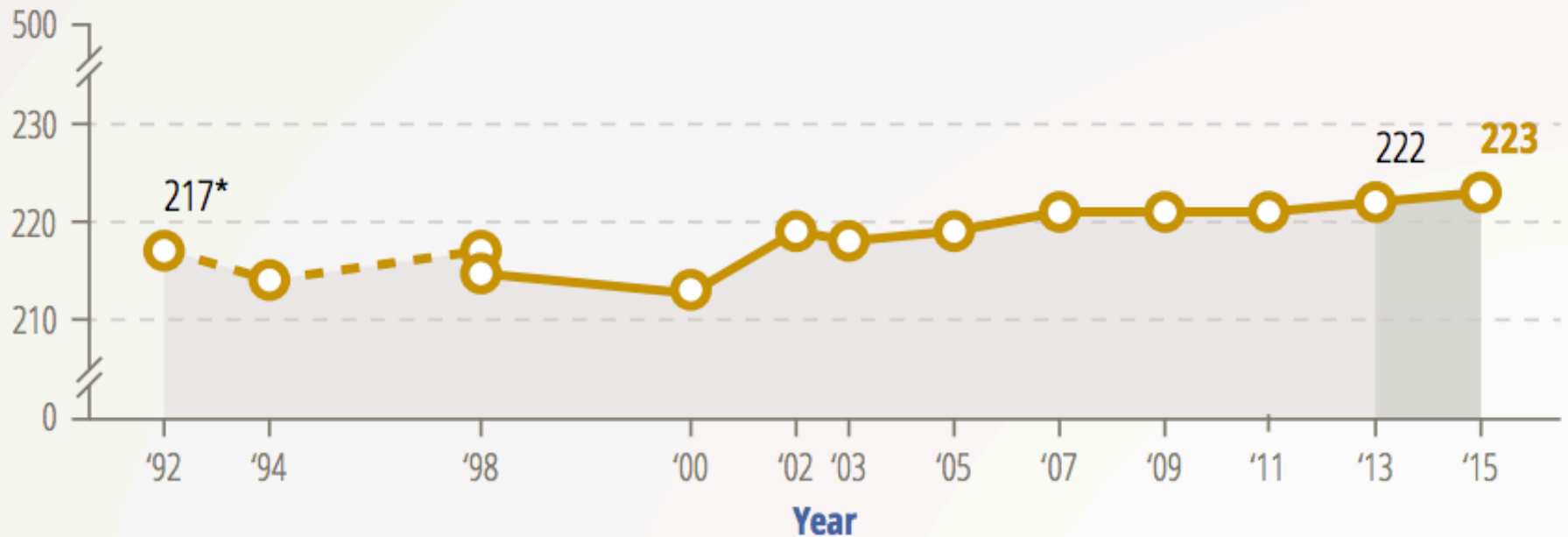
## Scale Score



\* Significantly different ( $p < .05$ ) from 2015

# Reading, grade 4

## Scale Score



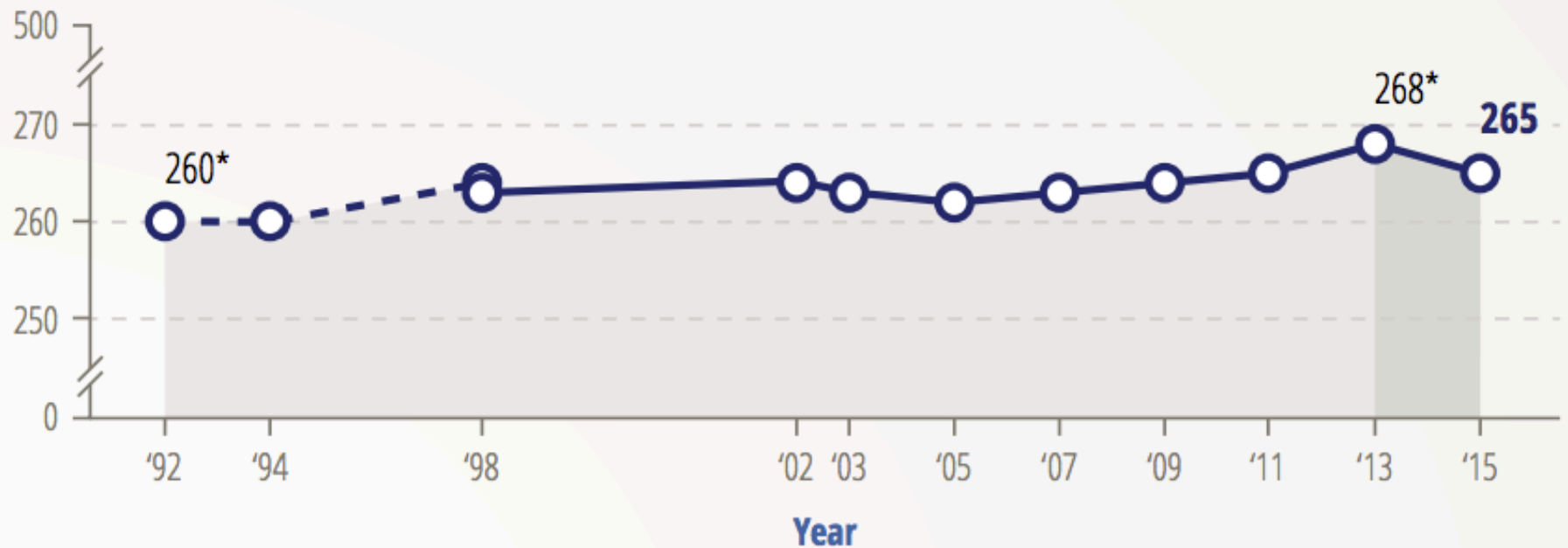
The dashed line indicates years in which NAEP was administered and accommodations were not available for students taking the assessment.

\* Significantly different ( $p < .05$ ) from 2015



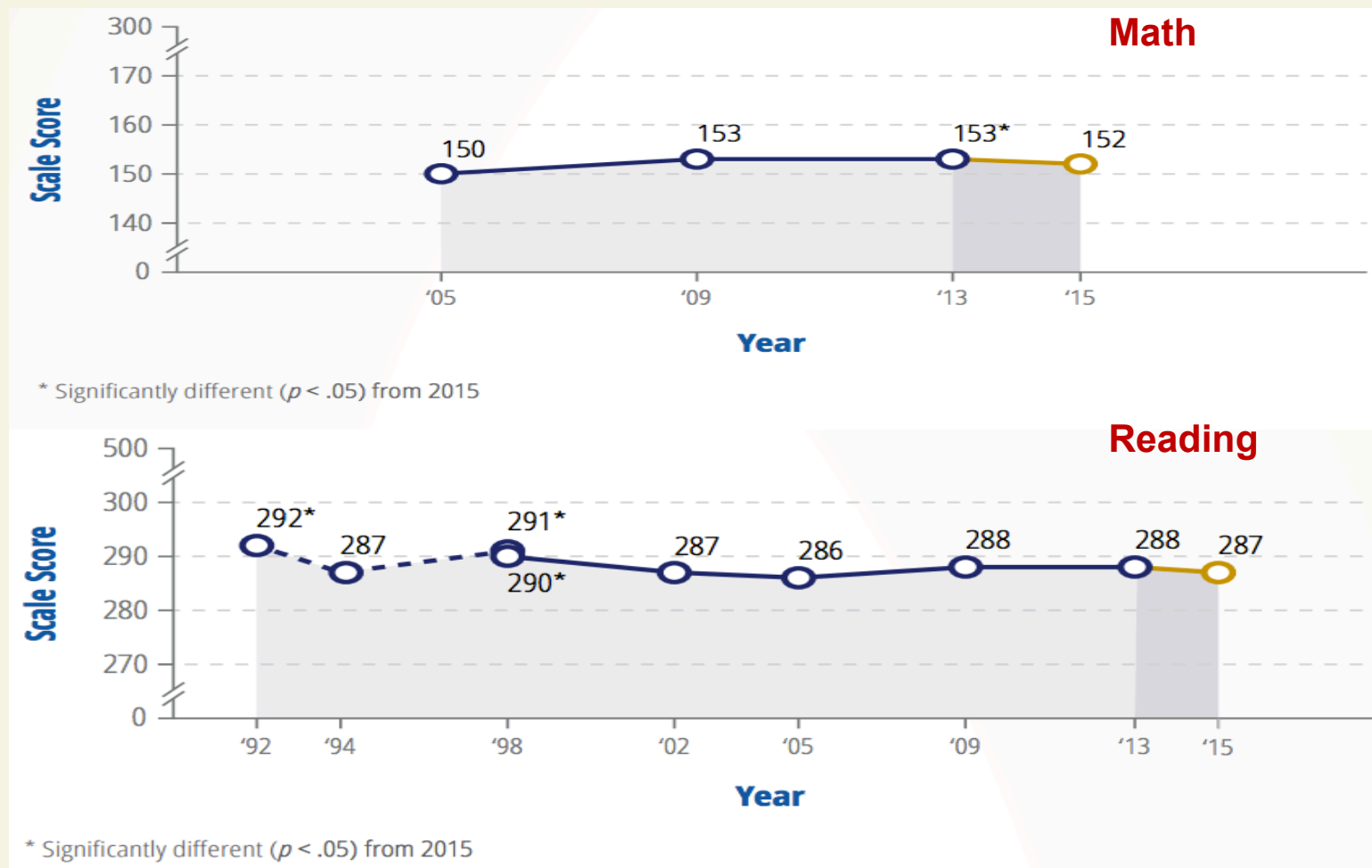
# Reading, grade 8

Scale Score



\* Significantly different ( $p < .05$ ) from 2015

# NAEP grade 12, math and reading



--- Dashed lines indicate years in which accommodations were not available for students taking the assessment.

NOTE: At grade 12, the NAEP mathematics scale ranges from 0 to 300, and the NAEP reading scale ranges from 0 to 500. Changes to the mathematics framework in 2005 necessitated starting a new trend line for that subject at grade 12.

# NAEP scores by achievement levels

