



Accountability Directors and Test Coordinators Fall Meeting

Accountability Services
North Carolina Department of
Public Instruction

September 16, 2015

READY



Mode Comparability Discussion Score Comparability: Paper-and-Pencil and Online

Kinge Mbella, Ph.D.
Psychometrician,
Accountability Services



Score Comparability: Paper-and-Pencil and Online

The focus of this discussion is to:

- Explain procedures used by Test Development to account for any variability
- Highlight key considerations when comparing scores from paper-and-pencil and online test administrations



Possible Sources of Score Variability

$$X_{ijk} = \mu \pm \alpha_j \pm \beta_k \pm \alpha\beta_{ijk} \pm e_{ijk}$$

- A goal when building assessments is to understand and be able to minimize systematic differences in student scores that are caused by:
 - Administering multiple forms α_j
 - Administering forms in two modes β_k
 - Administering multiple forms between modes to different students $\alpha\beta_{ijk}$

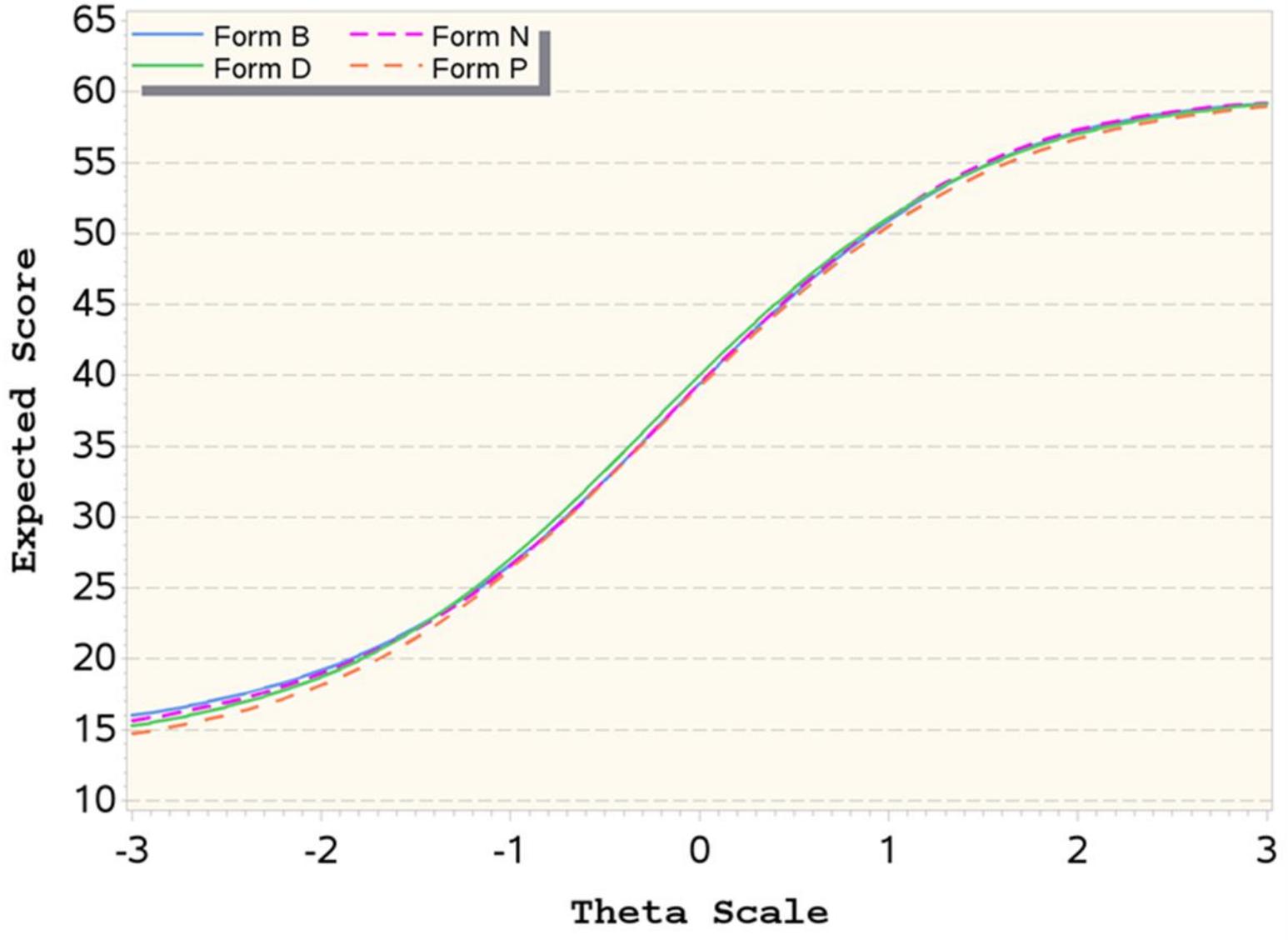


How do we ensure scores are comparable across forms and between modes?

- In terms of form building, statistical procedures are used to ensure forms are parallel.
- All parallel forms are built to match the same content and test specification.
- Each online form is the exact replicate of its paper counterpart with the exception of Technological Enhanced (TE) items in some EOC and EOG assessments.
- However, separate raw-to-scale tables are constructed for each form to control for any unintended differences in form difficulty not directly related to differences in student performance.



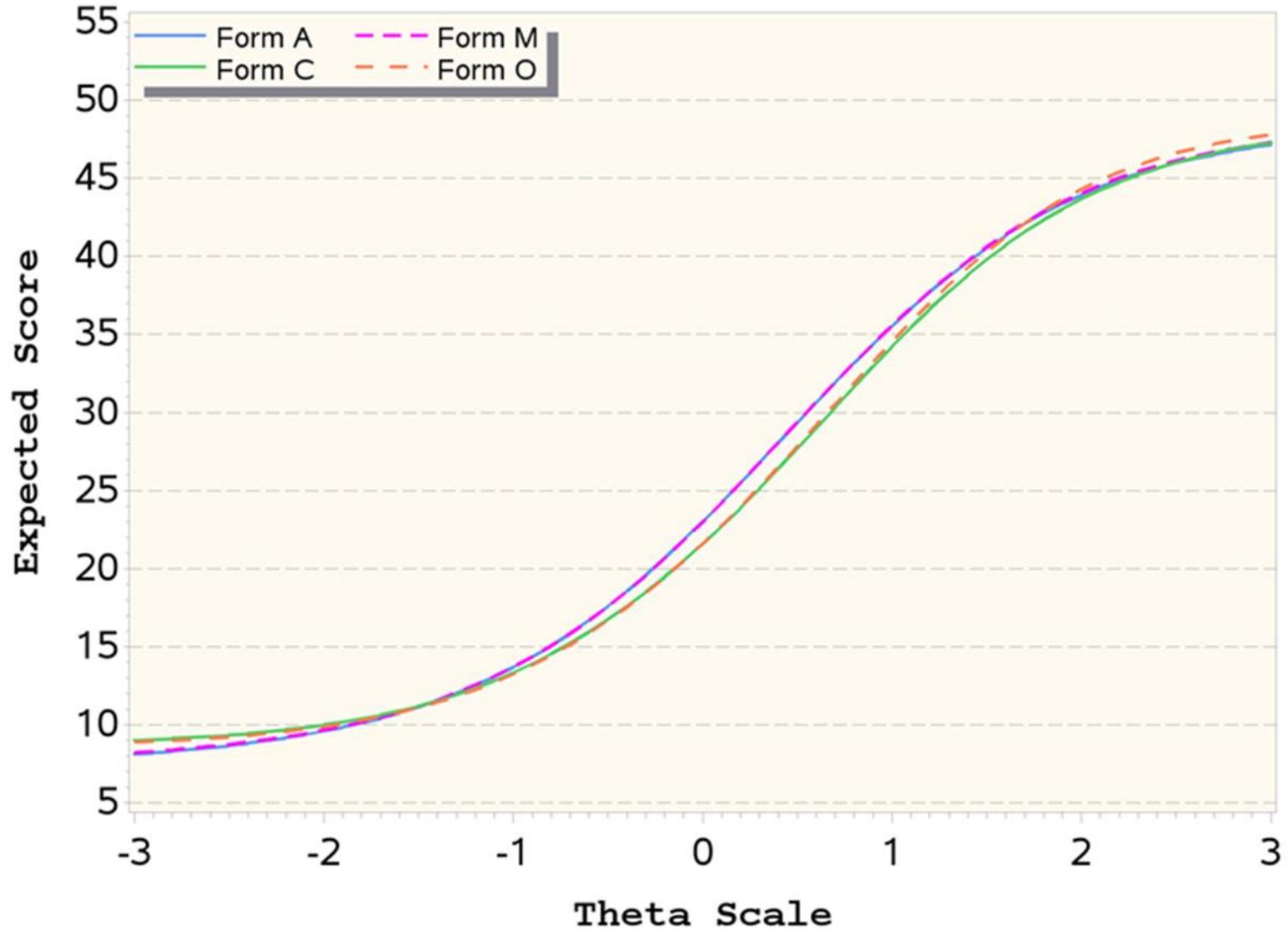
BIOLOGY Test Characteristics Curves of Operational Base Forms



Note: Forms B and D are paper forms. Forms N and P are online forms.



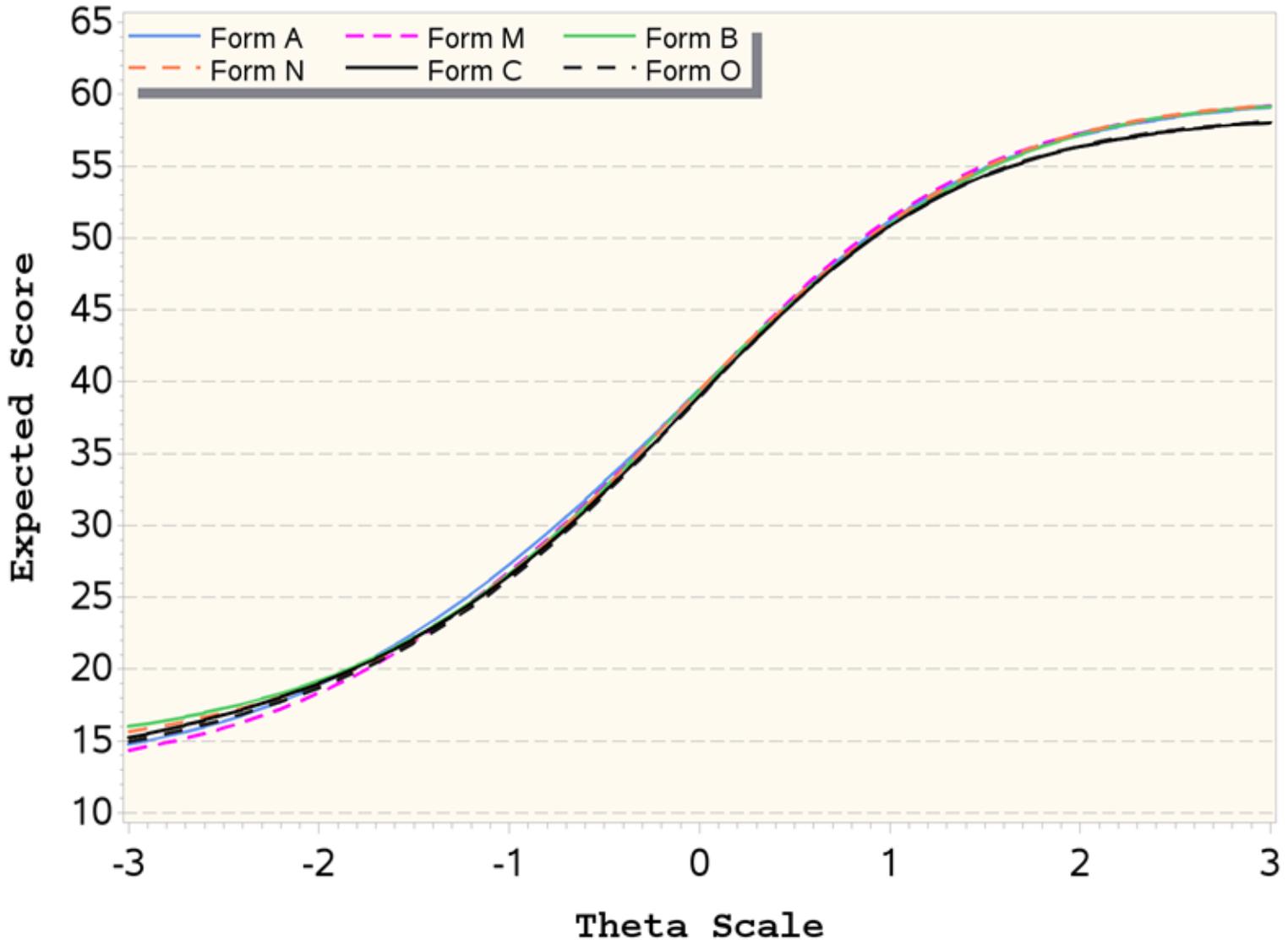
Math I Test Characteristics Curves of Operational Base Forms



Note: Forms A and C are paper forms. Forms M and O are online forms.



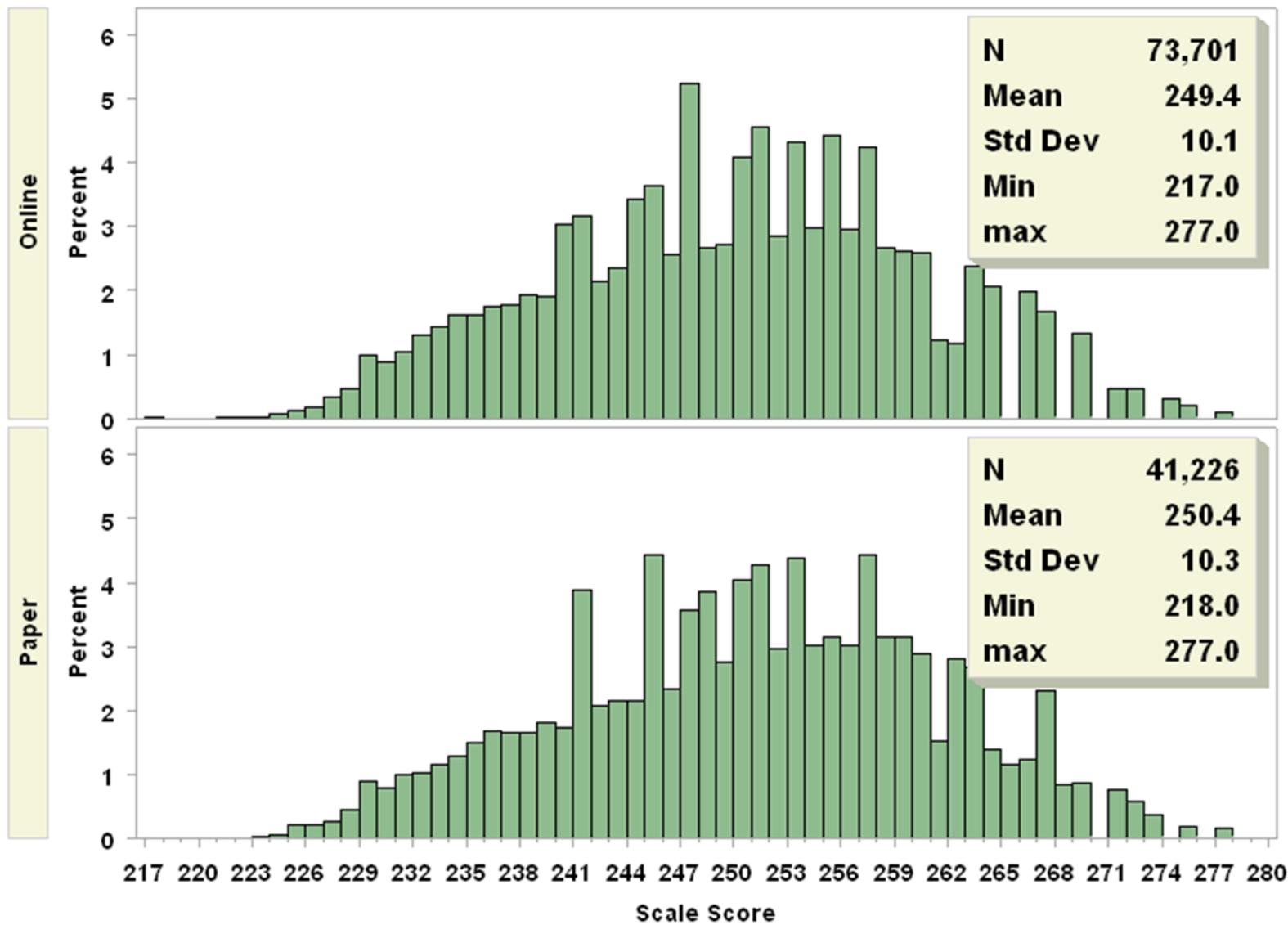
ENGLISH II Test Characteristics Curves of Operational Base Forms



Note: Forms A, B, and C are paper forms. Forms M, N, and O are online forms.

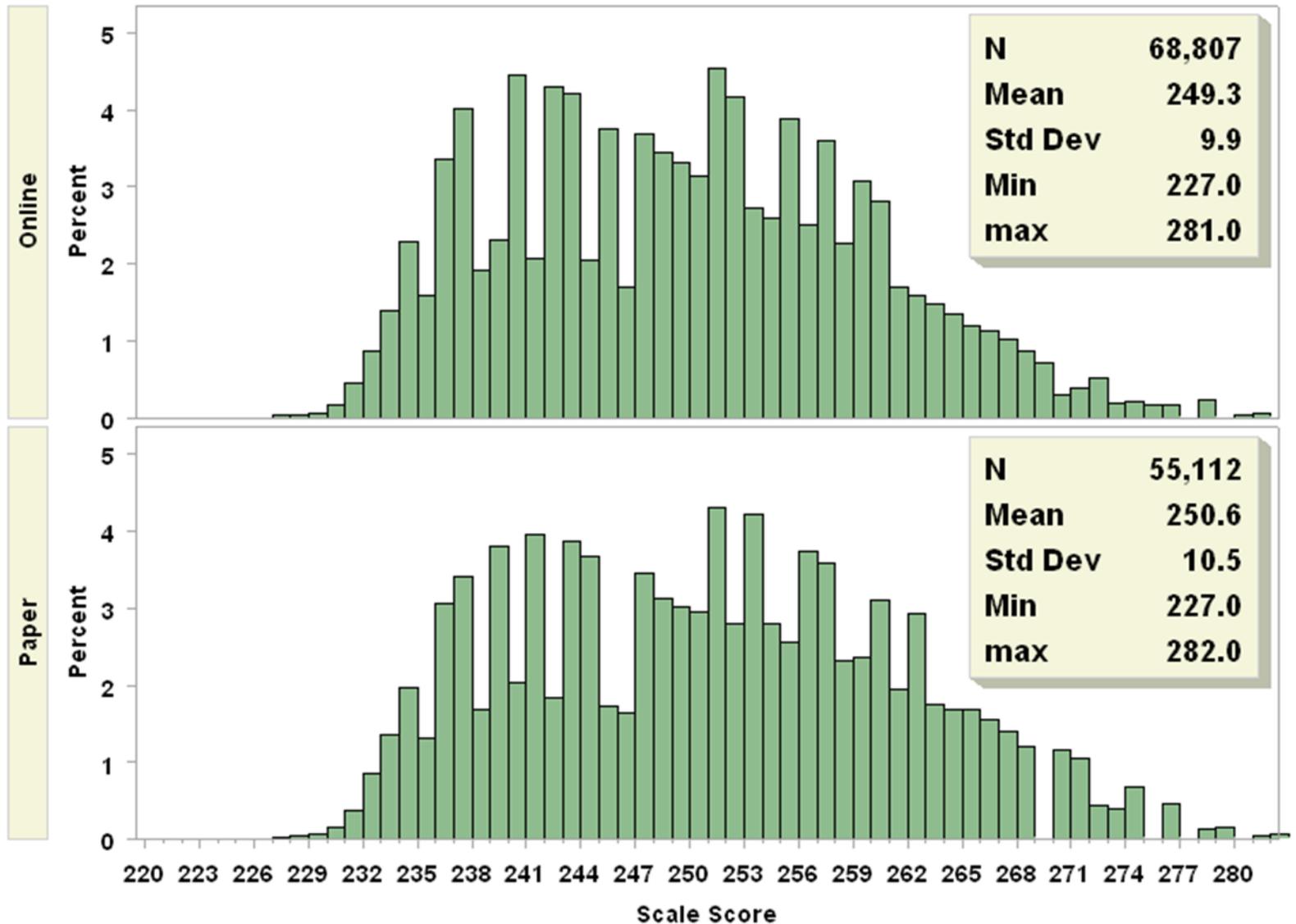


Biology Scale Score Comparison By Test Mode



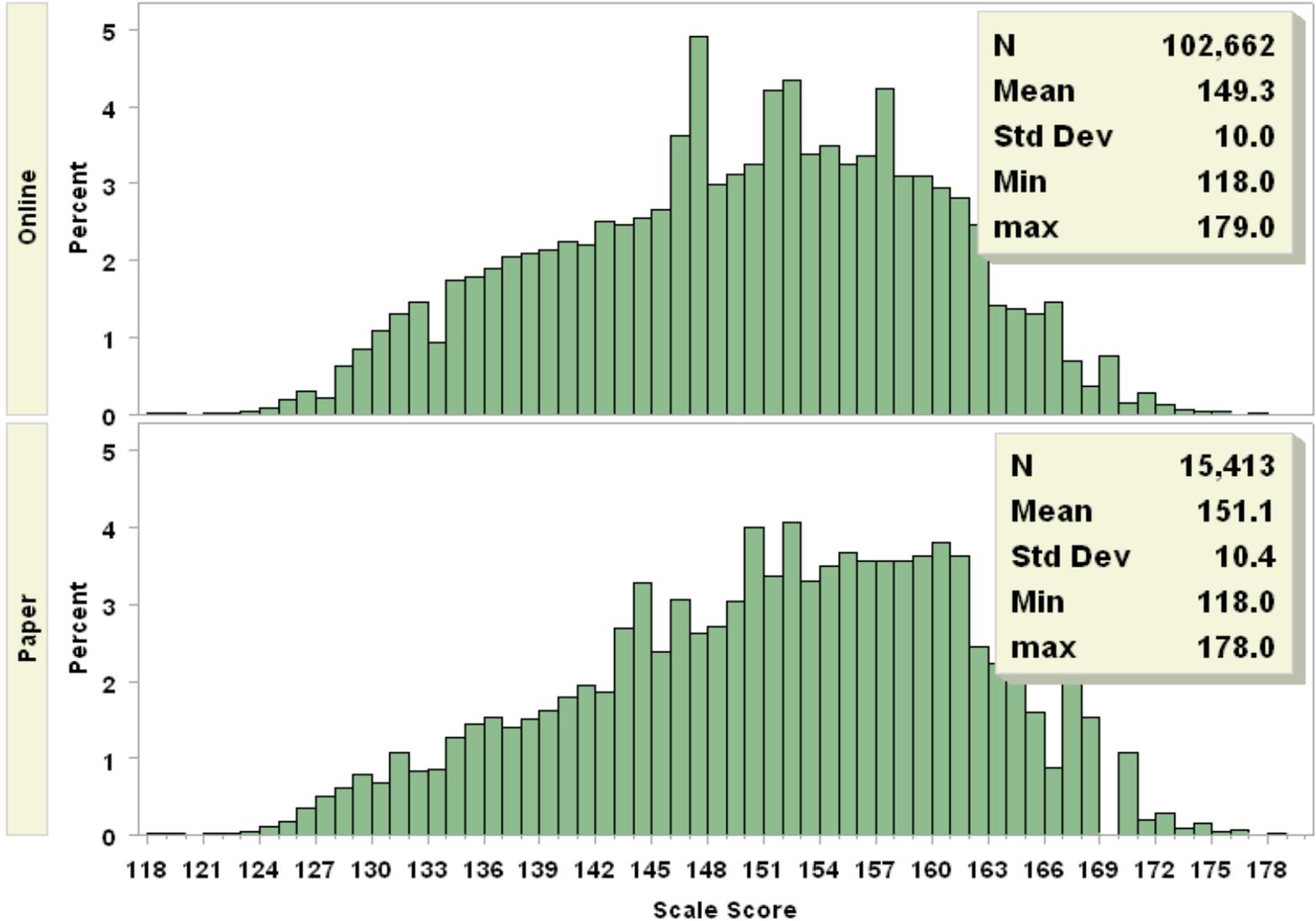


Math I Scale Score Comparison By Test Mode





English II Scale Score Comparison By Test Mode





Important Considerations on Score Comparability

- There are several factors that may confound score differences when doing comparability analyses
 - Generalizability- representativeness of sample
 - Interaction effect
 - Grade level



Caution–Comparability

If students are not randomly assigned to test modes the following precautions should be taken prior to reaching any conclusions:

- It is important to consider individual group differences and local assignment rules before comparing scores across modes.
- If there is evidence of a significant interaction effect between students and mode then any interpretation of mode main effect is not generalizable.

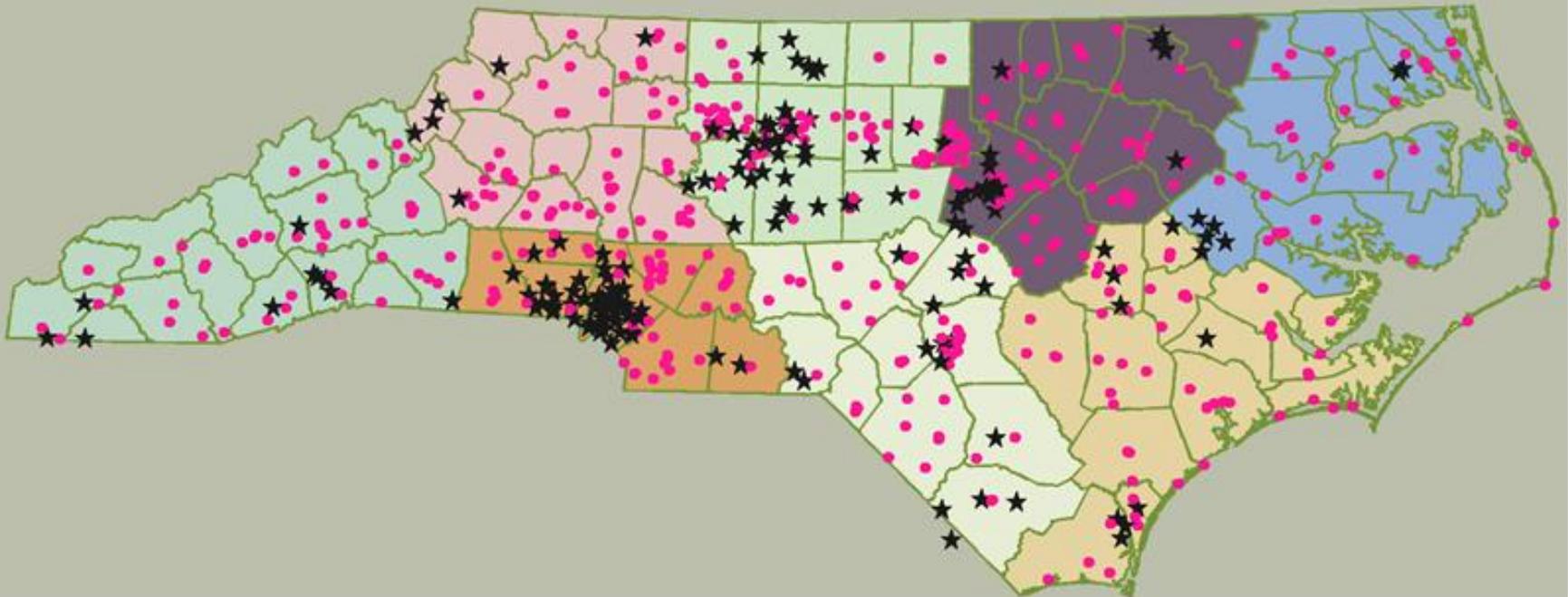


Sampling Effect: Test Mode Administered by Schools

Biology By Test Administration Mode

● = Online Forms

★ = Paper Forms

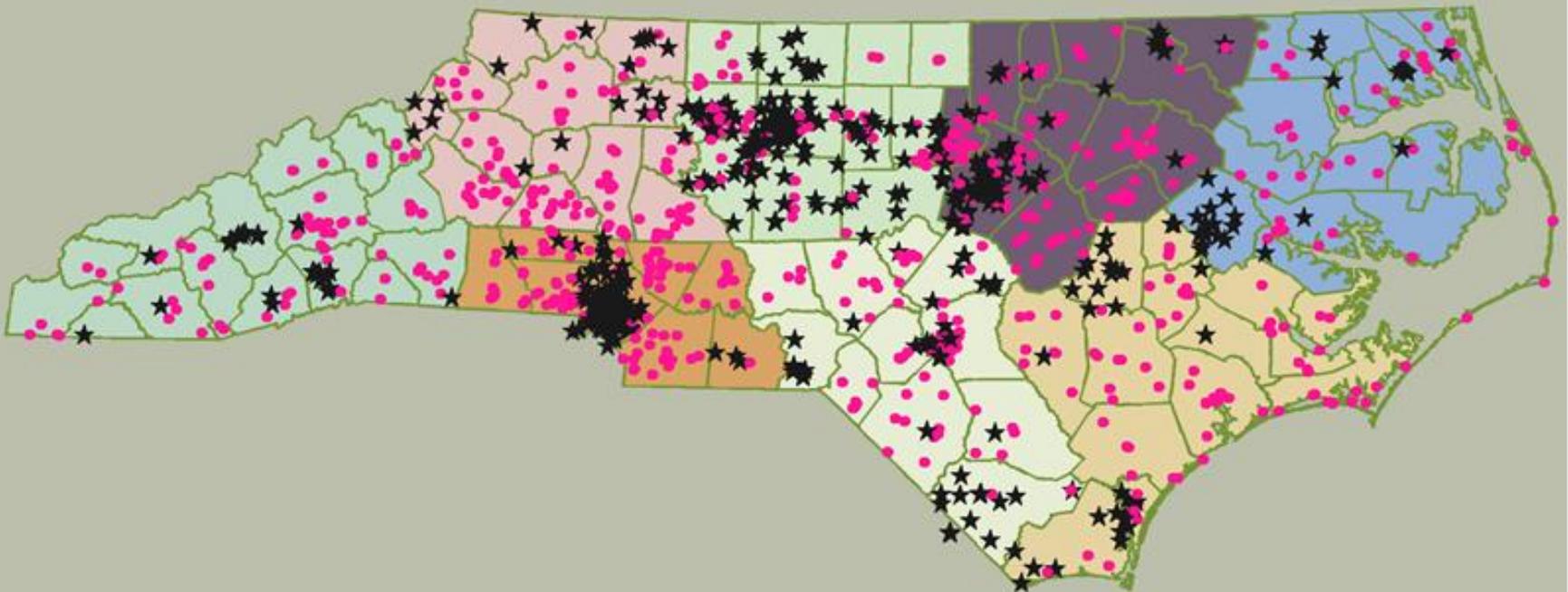




Sampling Effect: Test Mode Administered by Schools

Math I By Test Administration Mode

- = Online Forms
- ★ = Paper Forms



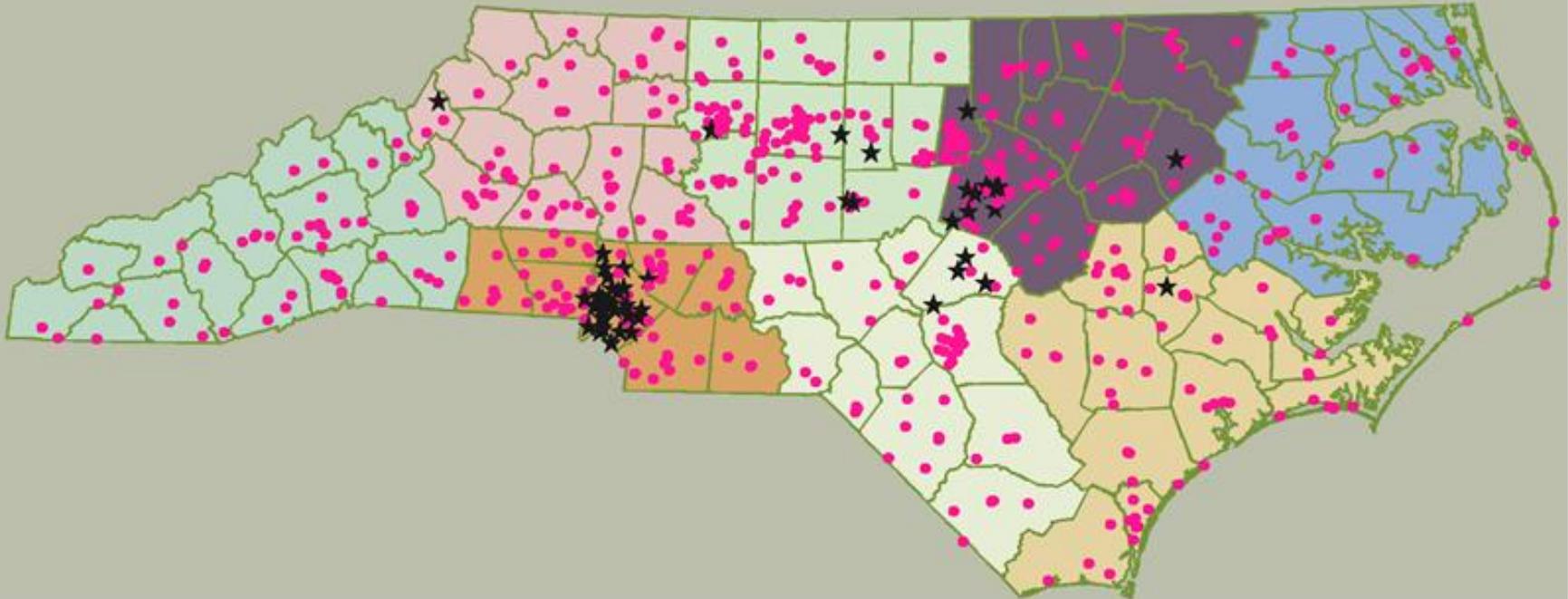


Sampling Effect: Test Mode Administered by Schools

English II By Test Administration Mode

● = Online Forms

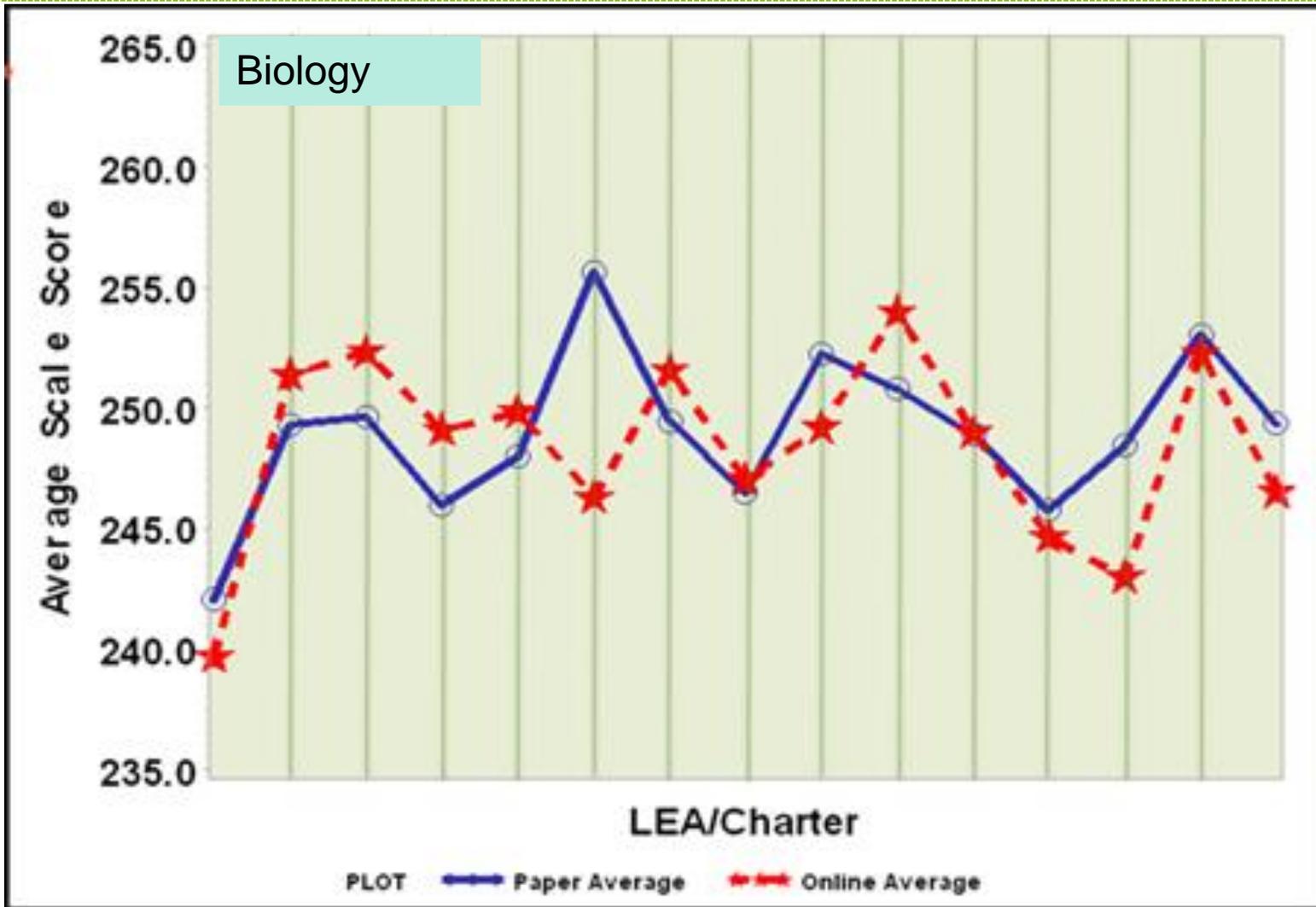
★ = Paper Forms





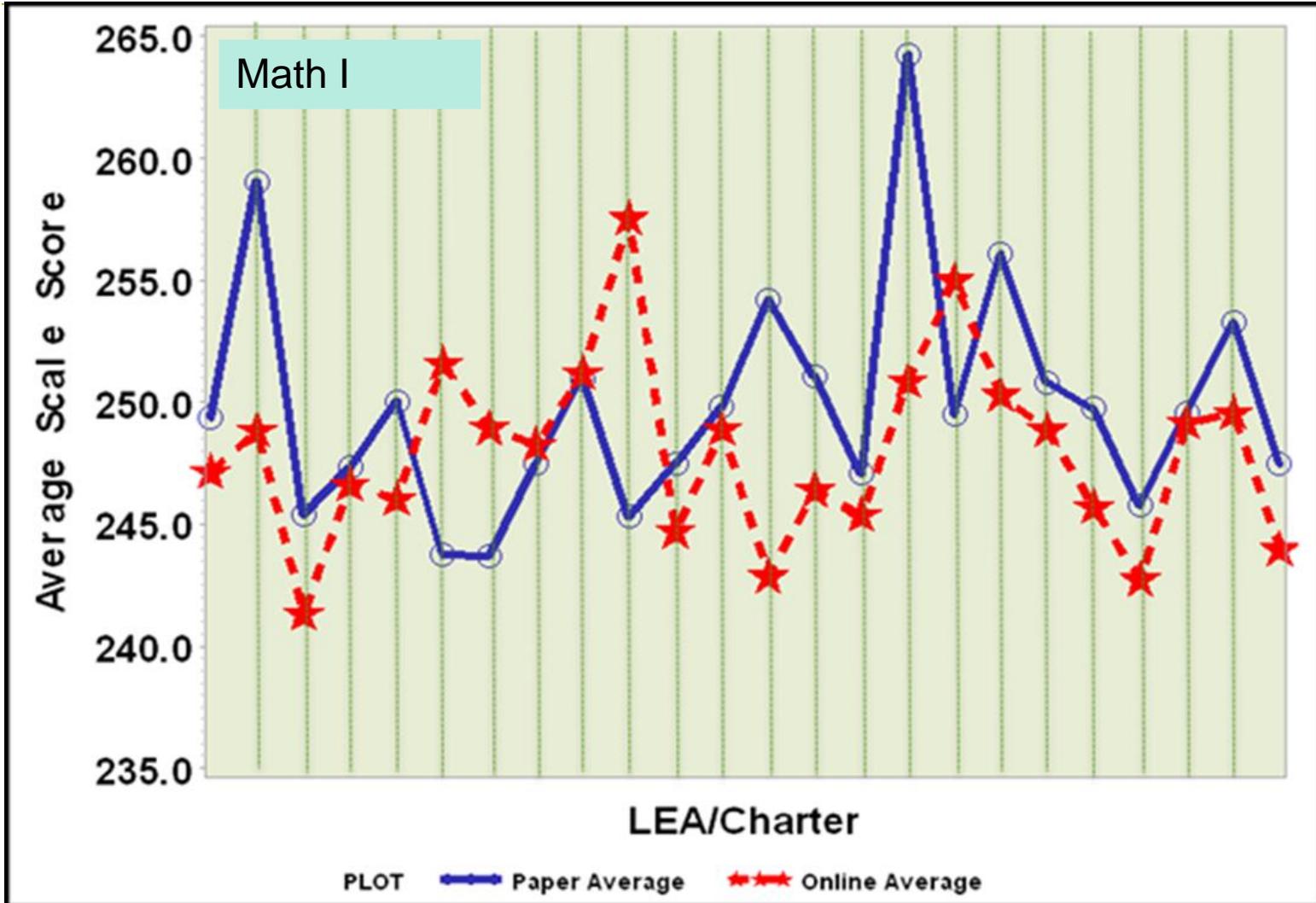
Interaction Effect: Biology EOC

Average Scale Scores for LEA/Charter Schools who Administered Online and Paper Forms





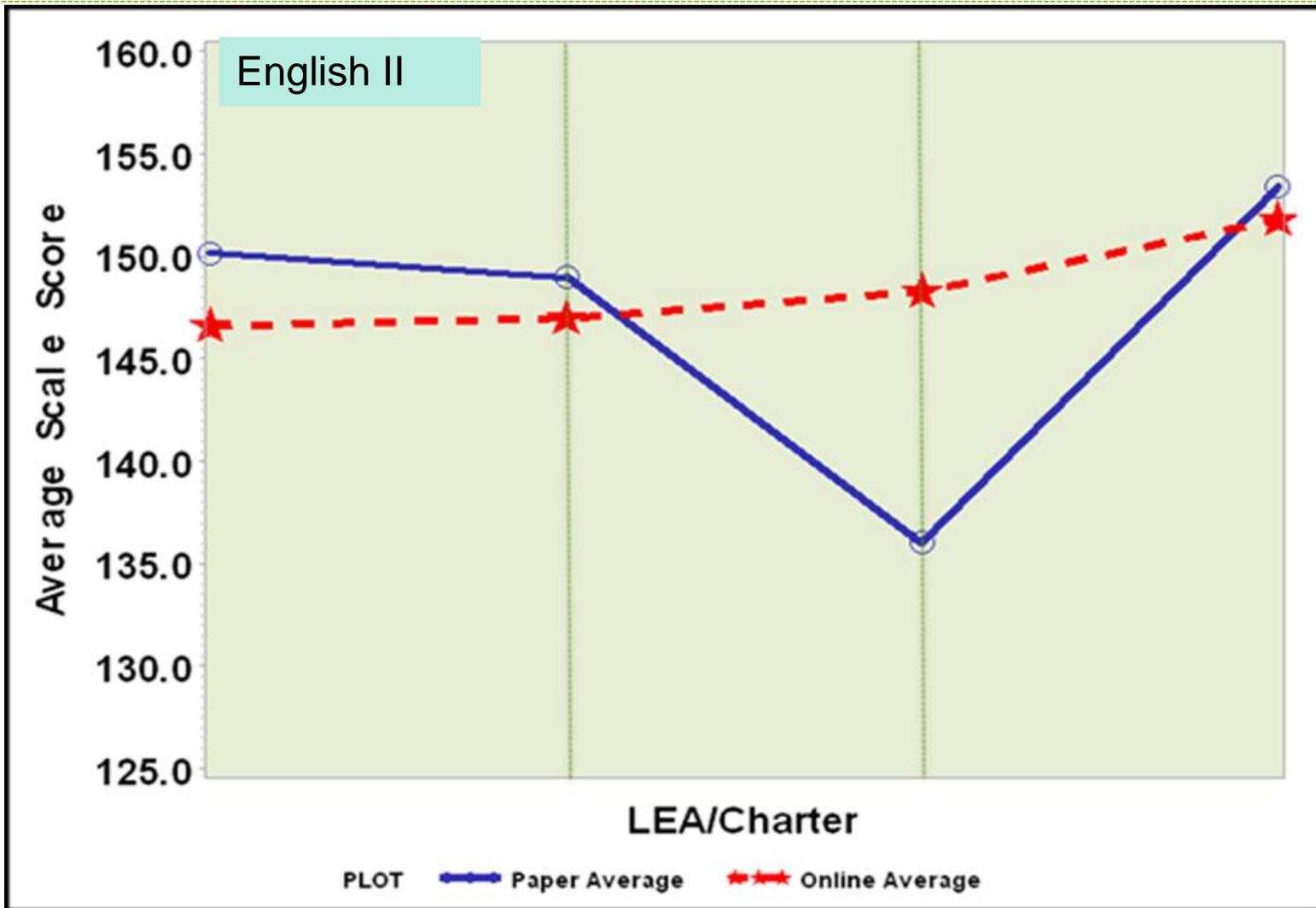
Interaction Effect: Math I EOC Average Scale Scores for LEA/Charter Schools who Administered Online and Paper Forms





Interaction Effect: English II EOC

Average Scale Scores for LEA/Charter Schools who Administered Online and Paper Forms





Audience Interaction



Example: EOC 2015



Poll:

Which Hypothesis is most Plausible?

Everything being equal:

- A. On average, students statewide perform significantly better on online forms.
- B. On average, students statewide perform significantly better on paper forms.
- C. On average, there is no significant difference on students' performance statewide between online and paper forms.



Everything being equal:

- A. On average, students statewide perform significantly better on online forms.
- B. On average, students statewide perform significantly better on paper forms.
- C. On average, there is no significant difference on students' performance statewide between online and paper forms.



Conclusion: Test Development Procedures Used to Account for any Variability

- All test forms for each EOC are constructed to be parallel or equivalent.
- Equating and scaling procedures are used to create separate raw-scale-score tables for each form.
- This ensures that differences in students' final score is not attributed to a particular form or mode of administration.

Note: When performing any comparability analyses across forms or between mode, it is imperative to make sure samples are matched and representative of students.



Comparability Questions to Discuss Locally

- Are online forms administered to different groups according to grade level (i.e. Math I for 8th graders)?
- Are online forms primarily administered to students who qualify for read aloud accommodation?
- What are the characteristics of schools that are administering online forms vs. paper forms within your district?
- How many student scores are being compared?



Online Assessment Review And Updates

Julie Molique
NC Department of Public Instruction

Sept 16, 2015

READY



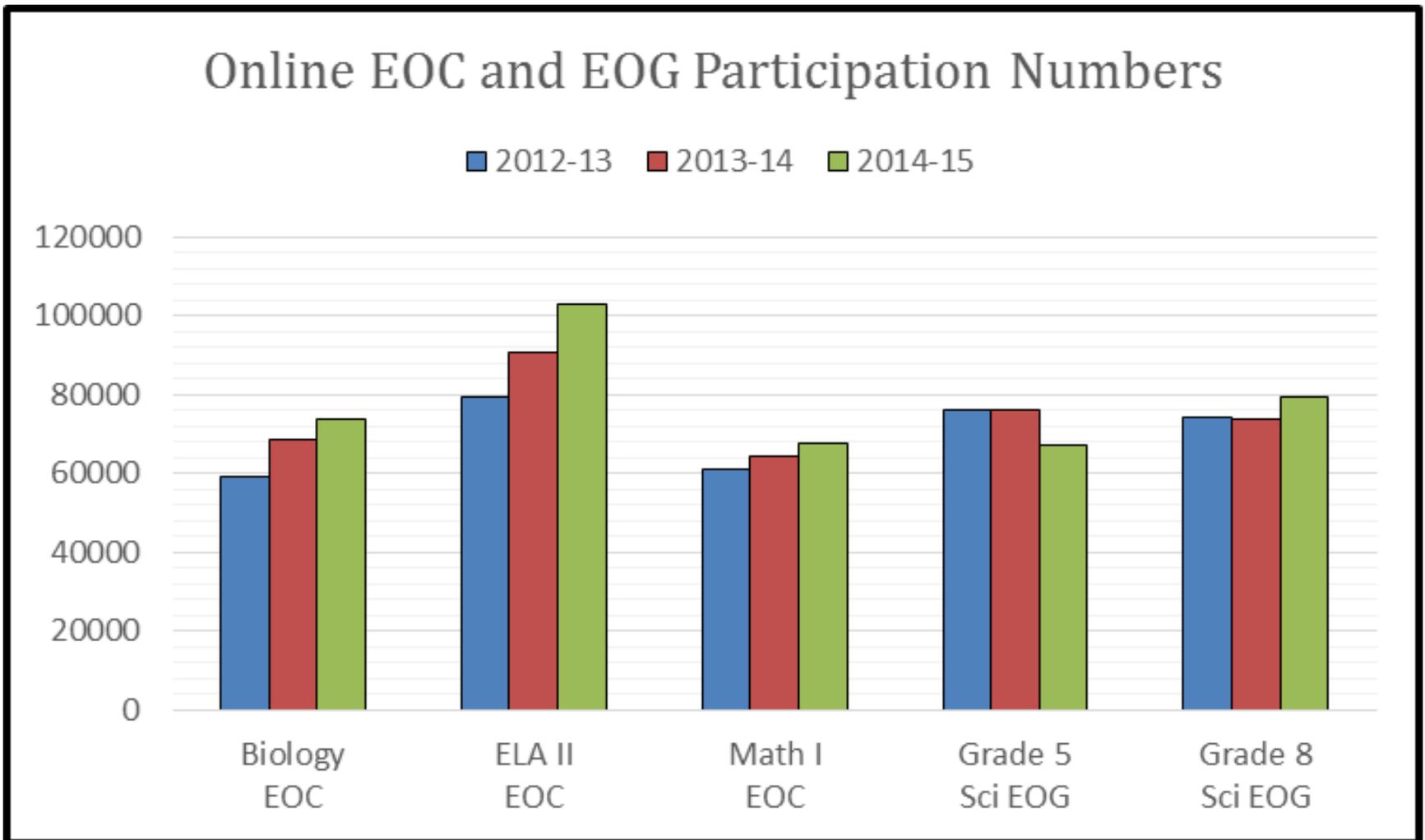
Online Participation Numbers

2014–15





EOCs and EOGs over the years.....





Percent Taken Online – EOC/EOG

