

Raise the Rigor: Designing Better Questions for Historical Thinking

Tra'Neisha LaRoche
Social Studies Subject Matter Expert

Many classroom assessments successfully measure what students remember but fall short of measuring how students think. As a result, students may perform well on recall-based quizzes yet struggle when asked to analyze historical evidence, justify conclusions, or evaluate historical significance. Designing rigorous assessment questions requires more than increasing difficulty; it requires intentionally matching the standard, Historical Thinking Skill, Depth of Knowledge, and assessment format to the thinking students are expected to demonstrate.

When it comes to assessment, rigor is not just about making something harder, it's about increasing the quality of thinking required to complete a task. In social studies, this means moving beyond questions that ask for recall and instead designing tasks that require students to analyze, justify, synthesize, and evaluate.

The EOCEP assesses student understanding using a range of item types, many of which operate at DOK 2 and 3. If our classroom questions never go beyond DOK 1, students will not be prepared to perform when the stakes are high.

Begin with the Standard

Effective assessment design begins before a question is ever written. First, identify the standard, the embedded Historical Thinking Skill, and the level of thinking students are expected to demonstrate. Once those expectations are clear, select the item type that provides students with an appropriate opportunity to show their understanding.

Understanding EOCEP Item Types:

- **Selected Response (SR):** Traditional multiple-choice with one correct answer.
- **Multi-Select:** Students choose more than one correct option.
- **Evidence-Based Selected Response:** Requires choosing a claim and then selecting the best evidence to support it.
- **Technology-Enhanced Constructed Response (TECR):** Interactive tasks like drag-and-drop or table completion based on sources.

Each type can be written at different DOK levels, but item structure should match what students will see on the EOCEP.

Matching Item Types to Historical Thinking

Selecting the appropriate item type is an instructional decision, not simply a formatting choice. Different Historical Thinking Skills often lend themselves to different types of evidence and assessment formats. For example:

- Comparison → Multi-Select or Evidence-Based Selected Response
- Causation → Constructed Response or Evidence-Based Selected Response
- Use of Evidence → Evidence-Based Selected Response
- Contextualization → Source-based Selected Response or TECR

The goal is not to force every skill into one item type but to choose the format that best allows students to demonstrate their historical thinking.

Increasing Cognitive Rigor with DOK:

- **DOK 1** asks students to recall or list facts.
- **DOK 2** requires understanding and using concepts (e.g., summarize, categorize, compare).
- **DOK 3** demands reasoning, justification, and multistep thinking.

Weak Task: DOK 1

Which New Deal program provided jobs to young men building parks and trails?

- A. CCC
- B. TVA
- C. SSA
- D. WPA

This is factual recall. It measures knowledge but not reasoning.

Better Task: DOK 2

Which two New Deal programs most directly reflect the government's expanded role in infrastructure and employment during the 1930s?

- A. TVA and SSA
- B. CCC and WPA
- C. WPA and SSA
- D. CCC and TVA

This version remains a multiple choice question but adds a conceptual layer by requiring students to connect programs to broader themes. Notice that the content did

not change, the thinking did. Raising rigor means increasing the reasoning students must use, not simply making the question more difficult. Although this example remains a selected-response question, the added reasoning better reflects the type of historical thinking students encounter on the EOCEP.

PLDs and Task Construction:

Use the PLDs to design assessment tasks that reflect different levels of student performance. For example:

- **Minimally Meets:** Match a New Deal program to its goal.
- **Meets:** Identify and explain the purpose of two programs.
- **Exceeds:** Use two sources to support a claim about how New Deal programs shifted public expectations of government.

Final Thought:

Raising the rigor means aligning the question type, DOK level, and performance level expectations. When students engage in rigorous tasks, they do more than prepare for a test. They learn to think like historians.

Effective assessment design begins with the standard and its embedded Historical Thinking Skill, matches the intended Depth of Knowledge, uses an appropriate EOCEP item type, and reflects the expectations described in the Performance Level Descriptors. When these elements work together, assessments do more than measure learning. They become tools for developing historical thinking.