

Sample #18: Essay Scoring Criteria, Secondary Social Studies

| Criterion | Indicator | Rating | Rationale (use words and phrases from the <i>Rubric for Rubrics</i>) |
|---------------------------|--|--------|---|
| Coverage/ Organization | 1A: Covers the Right Content | 5 | <ul style="list-style-type: none"> In the original report published by the authors,¹ essays by experts and novices were compared to discover the indicators of “deep understanding of subject matter content” (p. 3). In other words, individuals with deep understanding know a lot of information, understand how the information goes together into concepts, have few misconceptions, and can use information to interpret, analyze, and argue. So, these indicators didn’t come out of thin air, even though face validity (the ring of truth) is a little absent in a couple of cases. (The authors say that the best essays receive low-to-moderate scores for Text, with high scores on the other scales [p. 13].) |
| | 1B: Criteria Are Well Organized | 3 | <ul style="list-style-type: none"> We have some questions about organization for use with students. Do <i>General Impression of Content Quality</i>, <i>Prior Knowledge</i>, <i>Number of Principles or Concepts</i>, and <i>Misconceptions</i> overlap? That is, are the latter four actually indicators of the first? If they aren’t, on what other basis would one decide overall content quality? Or, it might be that the <i>Number of Principles or Concepts</i> is an indicator of the quality of argumentation, and therefore it might be profitable to combine these. Do <i>Prior Knowledge</i> and <i>Proportion of Text Detail</i> contradict each other? Would the nature of deep understanding be better communicated to students if indicators were recombined into two traits—knowledge and argumentation? |
| | 1C: Number of Levels Fits Targets and Uses | 5 | <ul style="list-style-type: none"> Five levels are generally within an appropriate range for secondary essays. |
| Clarity | 2A: Levels Defined Well | 2 | <ul style="list-style-type: none"> No levels are defined except for counts or vague comparison words like “0 = no response; 5 = highest level of understanding.” For the indicators listed, it seems that quantity is not a substitute for quality. For example, the use of two highly appropriate principles/concepts are probably a better indicator of deep understanding than four or more somewhat appropriate principles/concepts. We also wondered how one defines levels of quality argumentation. What specific indicators of quality argumentation might be useful to spell out for students? If students knew these criteria in advance, could they make sure that “seven or more pieces of information from the debates” and “four or more principles/concepts” are present in the response, regardless of how tangential? In other words, if students were to practice with these criteria, would their understanding deepen or would they merely find a way to achieve a high score? |
| | 2B: Levels Parallel | ?? | <ul style="list-style-type: none"> Hard to rate when levels are based on counts. |

¹ CRESST Performance Assessment Models: Assessing Content Area Explanations, CSE Report 652. (April 1992.) www.cse.ucla/products/reports_set.htm.