

# Rubrics on Historical Thinking Skills for Assignments and Student Work: Initial Validity Evidence

Emi Iwatani, Barbara Means, John Seylar, and Angela Hardy

January 15, 2021

## Suggested Citation

Iwatani, E., Means, B., Seylar, J., and Hardy, A. (2021). *Rubrics on historical thinking skills for assignments and student work: Initial validity evidence*. Digital Promise.  
<https://doi.org/10.51388/20.500.12265/112>

## Acknowledgements

This research was made possible through funding from Gates Ventures. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the funders.

## Contact Information

Email: [eiwatani@digitalpromise.org](mailto:eiwatani@digitalpromise.org)

Digital Promise:

**Washington, DC:**

1001 Connecticut Avenue NW, Suite 935  
Washington, DC 20036

**San Mateo, CA:**

2955 Campus Dr. Suite 110  
San Mateo, CA 94403

Website: <https://digitalpromise.org/>

## Executive Summary

Over the course of the year, Digital Promise sought to create and validate Historical Thinking Skills rubrics for use in its evaluation of Gates Ventures' World History Project (WHP) curriculum. Adopting a principled assessment development approach called *Evidence Centered Design* (Mislevy et al., 2003<sup>1</sup>), the Digital Promise team began by conducting an academic literature review in order to create an inventory of widely-recognized learning activities and historical skills. The results of this literature review (described in Hardy & Iwatani, 2021<sup>2</sup>) led to the creation of two sets of historical thinking skills rubrics, one for evaluating teacher-assigned activities (e.g., an essay prompt) and another for evaluating the student work those activities produced (e.g., a written essay). Using these rubrics, it was possible for a scorer to review teacher activities and student work to identify up to six historical thinking skills (called "dimensions"), and then assign a rating between 0 and 3 (called "progressions" or "scores") for each skill, depending on how advanced the activity or student work was along that dimension.

During early summer 2020, the Digital Promise team invited three expert World History teachers to apply this process to a collection of activities and student work. The objective of this scoring session, or Rubric Pilot Session, was to test the validity of two claims related to the historical thinking skills rubrics: (1) the rubrics could measure valuable historical thinking skills in terms of learning opportunities and outcomes and (2) the rubrics can be used consistently across trained scorers.

Analysis of the scores and scorer feedback from the Rubric Pilot Session supported both claims. First, the scores that trained raters assigned to activities and student work were generally consistent across scorers. Second, scorer feedback indicated that the rubrics did indeed measure valuable historical thinking skills. As WHP-sourced assignments were among the activities that the scorers reviewed, the Digital Promise team expected those assignments to score higher on the rubrics because of their focus on historical thinking skills. This expectation was met, providing further evidence that the rubrics are functioning in the ways they were intended.

In addition to validating these claims, the Rubric Pilot Session helped Digital Promise make needed revisions to both the Activity and Student Work rubrics (final versions presented in Iwatani et al., 2021<sup>3</sup>), substantially revise the scorer training protocol, and clarify the assignment collection process. These alterations serve to potentially increase the usability of the rubrics and make them suitable for the Digital Promise team's upcoming evaluation of the World History Project curriculum. These rubrics will be used at the end of SY 2021-2022 to support the evaluation of WHP's impact on students' historical thinking skills and students' opportunities to engage in historical inquiry.

---

<sup>1</sup> Mislevy, R. J., Almond, R.G., & Lukas, J.F. (2003). *A brief introduction to evidence-centered design* (Research Report 03-16). Educational Testing Service.

<sup>2</sup> Hardy, A., Iwatani, E. (2021). *Rubrics for examining historical thinking skills in high school world history activities and student work: Construct validity evidence from the literature*. Digital Promise.

<sup>3</sup> Iwatani, E., Hardy, A., Means, B., & Seylar, J. (2021) Rubrics for examining historical thinking skills in high school world history activities and student work. Digital Promise.

## I. Introduction

Digital Promise developed two sets of rubrics to support the evaluation of the World History Project: 1) “assignment rubrics” that assesses the extent to which classroom assignments provide opportunities for historical thinking, and 2) “student work rubrics” that assess the extent to which students are thinking historically in response to these assignments.

In educational measurement, validity refers to the degree to which evidence and theory support the interpretations of assessment scores for the proposed uses. Contemporary views of validity conceptualize it as an ongoing process that is initiated at the beginning of assessment design and continues throughout development and implementation. **This memo summarizes initial evidence (based on our rubric design and pilot), with respect to two validity claims:**

- **The rubrics measure valuable historical thinking skills (learning opportunities & outcomes)**
- **The rubrics can be used consistently across trained scorers**

It also describes how we developed, piloted and revised the rubrics.

## II. Rubric development process

To help us make sure that there are clear connections between what we observe in the assignments and what we infer about them, we used an evidence-centered design approach (Mislevy, Almond, & Lukas, 2003) to develop the rubrics (details in Appendix A). The development process included the following steps:

1. Identify the claims we want to make
2. Review literature on conceptions of historical thinking skills
3. Establish assessment targets (i.e., the dimensions of historical thinking skills)
4. Review literature on progressions for historical thinking skills dimensions
5. Draft rubric progressions
6. Revise draft rubric progressions for usability
7. Pilot draft rubric and solicit scorer feedback
8. Analyze rubric scores and scorer feedback
9. Revise draft rubric to final form

The draft rubrics were piloted in July 2020 by having three expert world history teachers score 19 world history assignments (e.g., worksheet or essay prompts) and associated 54 samples of student work (e.g., completed worksheets and essays) collected in Spring 2020. Digital Promise trained the scorers on the rubrics and analyzed their scores. In addition to scoring the sample assignments and student work, the scorers provided oral and written feedback on the usability of the rubrics and adequacy of the scoring process. Digital Promise reviewed both scores and scorer feedback for initial validity evidence and revised the rubrics accordingly. The profiles of the three scorers are provided in Table 1. The assignment collection process and the characteristics of the assignments are described in Appendix B.

Table 1. Scorer profile

Scorer 1 had 5 years of experience teaching World History and AP World History (with most of his experience in the non-AP course). He has worked within his school district to create and develop lesson studies and assessments that support new teachers. Among his priorities as a history teacher are a focus on non-Eurocentric primary sources and historical events for bias, reliability and perspective for his mostly-black students.
Scorer 2 had eight years of experience teaching in a public school settings, and has taught remedial, on-level and accelerated World History courses. With a B.A. in History and Latino Studies and M.A. in History, he has served as a department co-chair, supported district-level world history curriculum planning, and developed inquiry-based lesson plans for the History-Social Science Framework.
Scorer 3 had four years of experience teaching 10th grade world history at a Title 1 high school, and 15 years of teaching overall. With a Ph.D. in History, he has also given lectures at a nearby state university which serves a similar student population.

### III. Key findings from the pilot study

By analyzing the pilot rubric scores and scorer feedback we found that:

- There was good consistency among the 3 scorers overall in applying the assignment rubrics (ICC = .74) and student work rubrics (ICC = .71) (Appendix E & F).
- Scorers had difficulty distinguishing the assignment and work scoring process from everyday grading of student work, and from AP scoring, at least initially. Training and calibration took longer than we expected, suggesting the need for stronger, more targeted training.
- Focus groups and survey results indicated that scorers found all rubric dimensions to be important.
- The scorer feedback had more implications for the scoring/training process than for the rubric design (Appendix G).
- However, scores on two of the dimensions, “Claims and evidence” and “Comparison,” were not very consistent across scorers. Further analysis suggested that the former is likely because teachers are not used to assessing claims and evidence together (AP separates these), and the latter is likely because scorers readily identify “potential” opportunities for students to draw comparisons, even when comparison is not an explicit focus of the lesson. We believe both of these sources of scorer variation can be addressed with stronger and more targeted training (Appendix H).
- One lesson (Recipe for a Revolution) was very inconsistently scored. Upon further investigation, we found that this was likely because of ambiguities in the last portion of the lesson, rather than the rubric or scorers. For the evaluation, we will pre-screen assignment questions and have scorers flag any questions/assignments that seem ambiguous or otherwise problematic.
- On average, most assignments and student work scored between a 0 and 1 on each rubric, which is what we expected based on prior experience with this type of research

and based on our observations of the samples submitted (Appendix C). As WHP-sourced assignments were among the activities that the scorers reviewed, the Digital Promise team expected those assignments to score higher on the rubrics as their focus is on historical thinking skills. This expectation was met, providing further evidence that the rubrics are functioning in the ways they were intended (Appendix D).

- Assignments consisting only of pictures or diagrams could not be scored with the rubrics unless they were accompanied by written or oral explanations.

#### IV. Revisions made

Based on the rubric pilot, we decided to make the following modifications:

- **Modify the “Claims and Evidence” rubrics** to be more explicit about how assignments and work samples are to be scored, and to be about historical argumentation (rather than “claims and evidence”) so that there is less invitation to conflation with the AP rubric, which assesses claims and evidence separately (Appendix I).
- **Modify the rest of the rubrics slightly** to clarify and address scorer feedback (Appendix I). The main modifications were to add “historical” to characterize the skills and to emphasize that the assignments must “explicitly” call for students to employ them.
- **Substantially revise the training protocol** to include practice on components of the rubric (e.g., brief and extended arguments) to establish shared vocabulary before we have scorers rate an activity or piece of student work (Appendix H).
- **Tweak the assignment collection process** to include only assignments that require individual students to express themselves through prose (written or oral).

#### V. Summary of validity evidence to date and next steps

How do we know that the rubrics measure valuable historical thinking skills (learning opportunities & outcomes)?

- Use of evidence centered design helps to make sure that there are clear connections between what we observe in the assignments and what we infer about them.
- Broad literature base and scorer reviews supports the appropriateness and comprehensiveness of the dimensions.
- Assignments that were expressly designed to support historical thinking skills (namely the 5 World History Project assignments) scored higher on the majority of learning opportunities rubrics (all but Sourcing).

How do we know that the rubric can be used consistently across raters?

- The interrater reliability indices were good or excellent overall.

The next steps are to revise our training protocol, collect assignments and student work samples for our evaluation and to apply these rubrics.

## Appendix A. Evidence-centered approach to designing the rubric

One reason we believe the rubric measures valuable historical thinking skills is that **we used a principled approach to developing it**. Called *Evidence Centered Design* (Mislevy, Almond, & Lukas, 2003), the approach helps us to make sure that there are clear connections between what we observe in the assignments and student work, and what we infer about them.

In accordance with Evidence Centered Design, we articulated the types of claims we wanted to make (e.g., “World History Project promotes historical thinking skills and opportunities to learn them”), and defined the assessment targets (e.g., students understand historical causation). We carefully considered what evidence would imply that these targets are met (e.g., proficiency in historical causation could be inferred if students providing an accurate and extended analysis of a historical cause, rather than a brief/inaccurate analysis), and also considered tasks that would bring about the evidence (e.g., collect representative samples of student work from their world history class).

Each decision in the development phase needed adequate justification. For example, to define the six assessment targets for historical thinking skills, we conducted a thorough literature review of how historical thinking skills were conceptualized by academics and educators. Table A1. describes how we applied evidence-centered design to developing these rubrics.

Table A1. How the evidence-centered design (ECD) process was applied to creating the rubrics.

Development phase	Decisions made	Rationale
1. Defining the domain and claims	We want to measure opportunities for historical thinking skills in U.S. high school world history classrooms, and the extent to which students are thinking historically in these classes. We want to be able to claim that assignments provide (or don't provide) opportunities for students to engage in historical thinking skills, and that students in the class were proficient (or not) in them.	Developing historical thinking skills is considered to be an important goal of high school world history education by academics and educators alike and is an explicit goal of the World History Project.
2. Defining assessment targets	<ul style="list-style-type: none"> <li>• Assignment provides opportunity to [or student can] engage in historical argumentation (claims, evidence, consideration of possible objections)</li> <li>• Assignment provides opportunity to [or student can] source historical documents</li> <li>• Assignment provides opportunity to [or student can] make claims about historical <ul style="list-style-type: none"> <li>○ causation</li> <li>○ change and continuity over time</li> <li>○ comparison</li> <li>○ contextualization</li> </ul> </li> </ul>	A thorough literature review and review of 4 leading high school history standards/ frameworks supported these assessment targets. As explained below, feedback from scorers refined our articulation of the first target.
3. Defining evidence required	For each of the targets, we created a 4-point rubric articulating what we would need to observe to be able to infer that the targets are met. Generally, we	This aligns with protocols used in the past for assignment and student work analyses (e.g.,

	decided that lesson targets would be met if it explicitly asked students to provide extended arguments, evidence or analysis of these targets. Student targets would be met if students provided extended arguments, evidence or analysis.	Wenzel et al, 2002; Joyce, Gitomer & Iaconangelo, 2018) as well as with the literature on historical thinking skills and learning sciences, and team knowledge and experience in teaching and education research.
4. Developing the task models	We will apply the 4-point rubric to assess a representative sample of summative and everyday assignments and student work that are used in on-level or pre-AP 9th/10th grade world history classes. 6 pieces of student work will be randomly sampled from each class and averaged to produce a score that represents the class mean.	A direct way to assess whether teachers provide aforementioned learning opportunities is to examine whether these exist in their actual lessons. Whether students exhibit these skills in their assignments is also a direct way to assess whether students possess these skills.



**Appendix B. Assignment collection process used for the pilot**

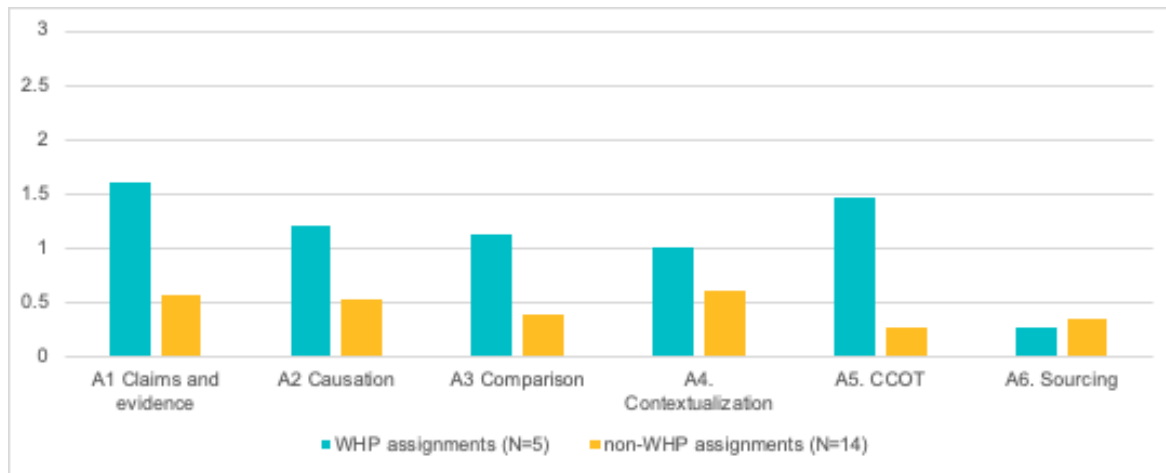
To collect the assignments used during the rubric pilot, 9th and 10th grade world history teachers were invited to submit examples of “everyday” or “summative” student work. Messages were sent to the online WHP community as well as three of Digital Promise’s relevant networks. Digital Promise supplemented these teacher submissions with a few publicly available online lessons from Facing History, the Stanford History Education Group, Choices, and WHP assignments available online (although too few of each to be able to make claims about these lessons). Among the 25 assignments Digital Promise collected through these channels, 4 were used for pilot scorer training and 19 were used for pilot scoring. Nine of these were summative assignments, which ranged from fairly common DBQ essay-style assignments to more novel assignments such as interviews with historical figures. Twelve assignments were formative or “everyday,” and could be expected to be assigned as a daily homework assignment or in-class activity. The classification of two assignments were unclear and could be used in either category. Scorers were instructed to score each assignment as well as up to five pieces of related student work if available - scorers simply scored a lesson if no student work was available.

Assignments and student work were de-identified and stored in a secure folder in Box. Scorers were provided with a spreadsheet populated with external links that allowed them to view the materials stored in Box.

### Appendix C. Descriptive statistics of rubric scores

<b>Rubric</b>	<b>N</b>	<b>Mean</b>	<b>SD</b>	<b>Min</b>	<b>Max</b>
<b>A1. Claims and evidence</b>	<b>19</b>	0.85	0.6	0	2.33
<b>A2. Causation</b>	<b>19</b>	0.72	0.82	0	3
<b>A3. Comparison</b>	<b>19</b>	0.59	0.54	0	1.33
<b>A4. Contextualization</b>	<b>19</b>	0.72	0.64	0	2
<b>A5. CCOT</b>	<b>19</b>	0.61	1.04	0	3
<b>A6. Sourcing</b>	<b>19</b>	0.33	0.72	0	2.67
<b>S1. Claims and evidence</b>	<b>12</b>	0.47	0.34	0	1.33
<b>S2. Causation</b>	<b>12</b>	0.63	0.76	0.06	2.67
<b>S3. Comparison</b>	<b>12</b>	0.38	0.48	0	1.67
<b>S4. Contextualization</b>	<b>12</b>	0.41	0.5	0	1.67
<b>S5. CCOT</b>	<b>12</b>	0.25	0.35	0	1
<b>S6. Sourcing</b>	<b>12</b>	0.06	0.13	0	0.44

## Appendix D. Rubric scores for WHP assignments compared to non-WHP assignments



Note: World History Project assignments included in the scoring were: Revolutionary Woman, Claim Testing - Globalization, CCOT - Transformation to responses, Marketing 101 - Forage or Farm, and Recipe for a Revolution. Each lesson had an explicit focus on one or more historical thinking skills. This is a small convenience sample of assignments from both WHP and the broader World History teaching community, and therefore are not necessarily representative of these two groups of assignments more generally. We were unable to make the analogous comparison for the student work rubric scores because only one of the World History Project assignments had student work associated with it.

**Appendix E. Consistency across raters (intraclass correlations) for assignment rubrics**

	<b>ICC1k</b>	<b>Qualitative rating*</b>
<b>A-series (114 scores/rater)</b>	.74	Good
<b>A1-Claims_and_Evidence only (19 scores/rater)</b>	.36	Poor
<b>A2-Causation only (19 scores/rater)</b>	.80	Excellent
<b>A3-Comparison only (19 scores/rater)</b>	.31	Poor
<b>A4-Contextualization only (19 scores/rater)</b>	.68	Good
<b>A5-CCOT only (19 scores/rater)</b>	.91	Excellent
<b>A6-Sourcing only (19 scores/rater)</b>	.87	Excellent
<b>A-series after dropping A1 only</b>	.79	Excellent
<b>A-series after dropping A3 only</b>	.79	Excellent
<b>A-series after dropping A1 &amp; A3</b>	.84	Excellent

\*Cicchetti (1994) provides commonly-cited cutoffs for qualitative ratings of agreement based on ICC values, with IRR being poor for ICC values less than .40, fair for values between .40 and .59, good for values between .60 and .74, and excellent for values between .75 and 1.0.

**Appendix F. Intraclass correlations for student work rubrics (student work averaged across assignments)**

	<b>ICC1k</b>	<b>Qualitative rating*</b>
<b>S-series (72 scores/rater)</b>	.71	Good
<b>S1-Claims_and_Evidence only (12 scores/rater)</b>	.03	Poor
<b>S2-Causation only (12 scores/rater)</b>	.92	Excellent
<b>S3-Comparison only (12 scores/rater)</b>	.48	Fair
<b>S4-Contextualization only (12 scores/rater)</b>	.65	Good
<b>S5-CCOT only (12 scores/rater)</b>	.63	Good
<b>S6-Sourcing only (12 scores/rater)</b>	.52	Fair
<b>S-series after dropping S1 only</b>	.76	Excellent
<b>S-series after dropping S3 only</b>	.75	Excellent
<b>S-series after dropping S1 &amp; S3</b>	.82	Excellent

\*Cicchetti (1994) provides commonly-cited cutoffs for qualitative ratings of agreement based on ICC values, with IRR being poor for ICC values less than .40, fair for values between .40 and .59, good for values between .60 and .74, and excellent for values between .75 and 1.0.

## Appendix G. Observations and revisions from the scoring process

### About the rubric dimensions and progressions (all dimensions)

- Scorers developed differing definitions of what constituted a “brief” or “extended” argument. We explained that this is determined by how well an argument is developed, rather than its length.
  - **Revisions:** Add this into training materials and FAQs
- In general, scorers commented that more examples can be useful to some teachers, including examples for Lv1 and 2
  - **Revisions:** Add examples to supplementary materials (training protocol & FAQ)
- Re: Usability - Reviews thought there needs to be clarification/explanation to how the rubric applies to non-traditional assignments and student work such as discussions, drawing or a media project (vs essays and short-responses).
  - **Revisions:** Clarification will be added to the sampling and training that individual students need to describe/explain (writing or speaking) regardless of what the medium is. If student work submitted do not contain such prose, then it should receive a low score.
- Description of each dimension only asks if there was an “opportunity” to demonstrate a skill
  - **Revisions:** Change language in each dimension to “explicitly call for” rather than “provides an opportunity to demonstrate” Title the dimensions to likewise match the rubric goals rather than the broader construct.
- Inconsistent scoring for Recipe for Revolution
  - **Revisions:** No change needed to the rubric itself (seems to have been more of an issue with the clarity of the assignment), but need to include a pre-screening process so the research team flags (and excludes) such lessons. Also request scorers to immediately notify facilitator of such (possibly) out-of-bound lessons.

### Claims and Evidence (A1 & S1)

There was general consensus on the importance of the dimension, clarity of its description and the ease of use. Specific feedback suggesting revisions/new development included:

- A1 (re: Levels) : Levels 0-3 not clear and distinct. Suggest to “add for a Level 1 that no evidence is used,” and suggest “having examples for every level; also make it clear what students are expected to include: is it a sentence at the end of introduction paragraph or is it throughout the essay?” (Rater 2)
  - **Revisions:** Revise rubric substantially to address this. Improve training so that this question (and others similar) are addressed and practiced by teachers.
- S1 (re: Wording): Some terms (e.g., “veracity”) may not be easily understood by students. (Rater 1)
  - **Revisions:** *No changes needed as this rubric is not intended for student use.*
- S1 (re: Helpfulness of student misconceptions): Seem unnecessary to include the example as most teachers would understand the difference between opinion and a thesis. (Rater 2)

- **Revisions:** *Retain the examples in case some teachers find it helpful.*
- S1 (re: usability by teachers to score their own students' work): "We had a discussion about grouping those two: claim and evidence together. I'm not sure that it's a good idea to do so though I see the importance of having made this choice--the claim should be grounded on historical evidence." (Rater 2)
  - **Revisions:** Revise rubric dimension substantially to address this.

Further review of scores found that two of the three scorers were consistently scoring lower on this dimension than intended, illuminating the need to strengthen the training for the scoring of this dimension, and to clarify the rubric wording so that scorers will not score a 0 if a either a student is asked to engage in historical argumentation (either by making, supporting or evaluating a historical claim).

### **Causation (A2 & S2)**

- A2 (re: usability for teachers to score their own lesson): "Causation is not an easy concept to understand for students and teachers; needs to be taught in an effective manner." (rater 2)
  - **Revisions:** Improve training so that this is explicitly addressed, and teachers "practice" identifying historical and non-historical causes; and extended and non-extended causal arguments in history.
- S2 (re: levels): "I think the score of "1" could be more clear if that needs to be historical. For the teacher lesson "The World in 2050" it was unclear how the rubric could be applied to a project like this."
  - **Revisions:** Change rubric to explicitly "Historical" mention in title and descriptors. Also add to Level 0 "OR Student employs causal reasoning in non-historical contexts"; Add to training that causal reasoning must be historical.
- S2 (re: Possible student misconceptions): "I'm wondering about scenarios where teachers asked students to "cook up" a recipe and students listed factors that led to the revolution. Does this kind of evidence count?"
  - **Revisions:** No revision necessary. Seems that scorer raises a clarification question about a particular assignment (vs provide feedback about the rubric).

### **Comparison (A3 & S3)**

- (no feedback specific to this category)

### **Contextualization (A4 & S4)**

- S4: (re: possible student misconceptions): "A common misconception might be whether a context just in the introduction is enough." (rater 2)
  - **Revisions:** In FAQ/supplemental resource - mention how this will not go beyond a 1. Also note that just mentioning a country, era or movement does not count (will be a 0 without any kind of "situating").

### **Continuity and Change Over time (A5 & S5)**

- A5 (re: description): "I would write out the acronym CCOT; it's not a common one used in the West Coast." (rater 2, (JS agrees))
  - **Revisions:** Write out CCOT everywhere

### Sourcing (A6 and S6)

- (no feedback on this category)

### About the scoring process and training

- Since the beginning of the scoring session, scorers had difficulty decoupling rubric scores and possible in-classroom implications, especially for student work. For example, scorers initially were reluctant to assign the low scores that the rubric required. This is because they were hesitant to give “good work” or “good lessons” low scores. Reiterating that there are many important skills that rubric does not touch (such as geography and historical empathy) and reminding scorers that the rubric does not get at lessons that are strictly about teaching content knowledge may help remedy this.
  - On a related note, it is important to reiterate that even “good” lessons and work should rarely, if ever, score 3s in all categories.
  - The in-classroom consequences that scorers were most apprehensive about were grades, and all three remarked that it seemed unfair to grade students harshly when they were not asked/prepared to demonstrate skills in the HTS rubric. For example, when a scorer scored so many categories with zeroes, they sometimes wanted to reward a marginally better category with a 3. Scores also expressed a desire to give students “benefit of doubt.” Training materials that explain that the HTS is not a grading tool (and not meant for students to ever look at) would aid in addressing this misconception.
  - **Revisions:** Modify training protocol to train for these aspects. Add practice sessions to clarify key definitions and distinctions before scorers score actual assignments and student work .
- Scorers remarked that the rubric could sound geared towards essay-type work. Including non-essay work examples could help with this.
  - **Revisions:** Train scorers on this specific point, and as suggested, include non-essay work samples in the training. In addition, tweak sampling protocol to make sure that student work is to include prose (i.e., cannot be non-verbal maps, figures, artwork).
- Scorers wondered whether assignments that “implicitly” promoted skills should merit higher scores. We explained that they should not.
  - **Revisions:** Train scorers on this specific point. Also revise assignment rubric to add that the skills should be explicitly called for.
- Some scorers found it odd that use of “claims” and “evidence” appeared on the same rubric, likely because they are separated on the AP rubric. In fact, one scorer explicitly asked why this rubric wasn’t more like the AP rubric. More effort should be taken to address this expectation and paradigm.
  - **Revisions:** Make substantial edits to the rubric dimensions themselves and have training that emphasizes key distinctions in these rubrics.
- One scorer remarked that the training would have been more helpful for him had it included more intentional “I do”, “we do” and “you do” statements. He feels that he would have been able to be more efficient.
  - **Revisions:** Include in the protocol



## Appendix H. Training Protocol Outline

### Training process:

- Explain the purpose of the scoring project and rubric.
- Before introducing any rubrics, introduce key terminology and distinctions that occur in the rubrics and train (quiz) scorers on these until scorers understand them fully. Such distinctions include: extended vs brief arguments; historical vs non-historical claims/causation/comparisons/changes/sources, making claims vs evaluating claims vs supporting claims; explicit vs implicit prompting; contextualization (how just mentioning a time and place in the beginning of an essay would not count).
- Introduce each rubric dimension for activities and student work and direct scorers to examples of assignments and student work that would score high in each dimension. Train the scorers to score sample assignments and student work by:
  - First, having the facilitator model the scoring for each dimension, explaining the rationale for the scores that they chose. Discuss scores as a group and address any inconsistencies or misconceptions.
  - Second, have scorers score a lesson together as a group. Discuss scores as a group and address any inconsistencies or misconceptions.
  - Finally, allowing scorers to score sample lessons and student work independently. Discuss scores as a group and address any inconsistencies or misconceptions. If scores are not consistent, continue to practice on training materials until consistency is established.
- Calibration items should be included in the scoring set, which should be monitored daily for any drifts in scoring. Re-calibration as necessary. Schedule regular check-ins and have lines of communication open between scorer and facilitator (e.g., office hours, email/slack channels) for scorers to be able to have their questions answered.
- Request reviewers to flag assignments or assignment sections that seem “out of bounds” -- e.g., a geography or study skills assignment; assignments with instructions that most students would likely have misunderstood.

### General Notes: (these can also go into an FAQ that accompanies the rubric)

- Remind scorers not to infer missing details about assignments from student work - When assignments lack helpful details about, for example, the scope of an assignment (is it formative or summative?), scorers should not use student work to try to infer what type of work the assignment is intended to produce. For example, it could be that a teacher is using a DBQ assignment that calls for an extended, detailed response as a warm-up, and that students are only expected to produce an essay outline or thesis. In this case, the assignment would score highly in A1, even though student work would not score as highly. It is important to refer to the teacher’s lesson plan to help avoid this pitfall.
- What is the difference between brief and extended arguments? - Almost every dimension uses this dichotomy. In student work, explain that it is not the length of the argument but the complexity of the argument that determines whether it is brief or extended. In assignments it can be more difficult to determine whether they call for a

brief or extended response, but the following factors can help: 1) whether the assignment is formative or summative, 2) how much evidence students are expected to employ 3) the time taken or length of the argument students are expected to produce.

- Decouple scores and grades - Scorers have had difficulty decoupling rubric scores and possible in-classroom implications, especially for student work. For this reason, scorers may be reluctant to assign the low scores that the rubric called for since they were hesitant to give “good work” or “good lessons” low scores. Reiterating that there are many important skills that rubric does not touch (such as geography and historical empathy) and reminding scorers that the rubric does not get at lessons that are strictly about teaching content knowledge may help remedy this.
  - On a related note, it is important to reiterate that even “good” lessons and work should rarely, if ever, score 3s in all categories.
  - The in-classroom consequences that scorers were most apprehensive about were grades, and all three remarked that it seemed unfair to grade students harshly when they were not asked/prepared to demonstrate skills in the HTS rubric. For example, when a scorer scored so many categories with zeroes, they sometimes wanted to reward a marginally better category with a 3. Scores also expressed a desire to give students “benefit of doubt.” Training materials that explain that the HTS is not a grading tool (and not meant for students to ever look at) would aid in addressing this misconception.
- Explicit mentions only - Scorers wondered whether assignments that “implicitly” promoted skills should merit higher scores (ex. “Students *could* use a comparison to make this argument...”). Explain that only explicitly required activities should count towards a score in the Activities rubric.
- Untrain a propensity toward AP scoring style - Some scorers found it odd that use of “claims” and “evidence” appeared on the same rubric, likely because they are separated on the AP rubric. Effort should be taken to address this expectation and paradigm.

**Appendix I. Pilot rubric with changes tracked**

## Rubrics for Examining Opportunities for Historical Thinking Practices in High School World History Activities

### Introduction

This set of rubrics are meant to be used to study the extent to which high school world history classroom activities provide opportunities for students to engage in historical thinking practices.

The activity dimensions A1 - A6 can be used to assess the extent to which a learning activity provides students the opportunity to:

- Make and develop claims and/or assess the quality of claims found in a historical account or interpretation (A1 - Historical argumentation Claims and evidence)
- Employ causal reasoning (A2 - Historical Causation)
- Describe and explain similarities and differences between historical developments, processes, regions, eras, or other focal areas (A3 - Historical Comparison)
- Contextualize historical phenomena and actions within a temporal, spatial and/or sociocultural setting (A4 - Historical Contextualization)
- Analyze continuity and change over time in history (A5 - Continuity and Change Over Time in History)
- Source a historical document (e.g., identify the author's purpose and perspective) (A6 - Sourcing)

Each dimension has four levels (0-3) where generally 0 indicates the absence of a historical thinking practice in the activity, 1 indicates emergence, 2 indicates partial presence, and 3 indicates solid presence.

**A1. HISTORICAL ARGUMENTATION CLAIMS AND EVIDENCE:** Activity **explicitly calls for students** ~~provides students the opportunity to make, support or assess~~ **make and develop a historical claim and/or assess the quality of a claim found in a historical account or interpretation.**

0	1	2	3	Level 3 Examples
Activity <b>does not explicitly</b> call for students to make, support or assess a historical claim- <del>claim supported by evidence and reasoning and/or assess the quality of a claim.</del>	The activity <b>explicitly prompts students to state</b> a historical claim, a reason for a historical claim or an evaluation of a historical claim <del>make a claim and/or make an evaluative statement about the quality and veracity of a claim in a historical account and/or interpretation.</del>	The activity <b>explicitly</b> prompts students to state a historical claim, a reason for a historical claim or an evaluation of a historical claim <del>make and develop a claim and/or evaluate the quality and veracity of a claim in a historical account and/or interpretation.</del>  <b>The activity explicitly prompts students to briefly explain their claim, reason, or evaluation</b> <del>make a brief argument using evidence and reasoning to support their own claim, and/or to support, qualify, or refute the validity of another person's claim.</del>	The activity <b>explicitly</b> prompts students to state a historical claim, a reason for a historical claim or an evaluation of a historical claim <del>make and develop a claim and/or evaluate the quality and veracity of a claim in a historical account and/or interpretation.</del>  <b>The activity explicitly prompts students to provide an extended explanation</b> of their claim, reason, or evaluation. <del>The activity prompts students to make an extended argument using evidence and reasoning to support their own claim, and/or to support, qualify, or refute the validity of another person's</del>	For example, the activity may prompt students to support their own claim by <ul style="list-style-type: none"> <li>• integrating relevant evidence from multiple sources,</li> <li>• acknowledging the credibility and limitations of the evidence used and noting discrepancies across sources,</li> <li>• addressing and evaluating potential counterarguments, and/or</li> <li>• using reasoning to connect the evidence to their claim.</li> </ul> Or, the activity may prompt students to extend their argument on the validity of another person's claim by <ul style="list-style-type: none"> <li>• examining supporting and refuting evidence, including information the claim maker has omitted,</li> <li>• explaining the strengths and limitations of the reasoning used to support the claim,</li> <li>• corroborating the claim with other accounts or interpretations, and/or</li> <li>• considering the perspective and credibility of the claim maker.</li> </ul>

			claim.	
--	--	--	--------	--

\*Historical claims can include historical accounts and interpretations.

**A2. HISTORICAL CAUSATION:** Activity **explicitly calls for students** ~~provides students the opportunity~~ to employ causal reasoning using appropriate historical evidence.

0	1	2	3	Level 3 Examples
The activity <b>does not explicitly</b> call for students to describe causes and/or effects to explain human actions, events, and/or larger structures or processes.	The activity <b>explicitly</b> calls for students to <b>describe causes and/or effects to explain human actions, events, and/or larger structures or processes.</b>	The activity <b>explicitly</b> calls for students to describe causes and/or effects to explain human actions, events, and/or larger structures or processes, <b>and asks students to provide a brief analysis of distinctions between different causes and/or effects (e.g., primary vs. secondary or immediate vs. long-term) or the relationship between causes and/or effects.</b>	The activity <b>explicitly</b> calls for students to describe causes and or effects to explain human actions, events, and/or larger structures or processes, and asks students to provide <b>an extended</b> analysis of distinctions between different causes and/or effects (e.g., primary vs. secondary or immediate vs. long-term) or the relationship between causes and/or effects.	For example, the activity may prompt students to <ul style="list-style-type: none"> <li>• provide an extended analysis of the distinctions between background conditions, triggering events, primary and secondary causes, and/or immediate and long-term effects.</li> <li>• evaluate the relative historical significance of various causes and effects.</li> </ul>

**A3. HISTORICAL COMPARISON:** Activity **explicitly calls for students** ~~provides students the opportunity~~ to describe and explain similarities and differences between historical developments and processes, regions, eras, or other focal areas, using appropriate historical evidence.

0	1	2	3	Level 3 Examples
Activity <b>does not explicitly</b> call for students to describe similarities and differences between historical developments, processes, regions, eras, or other focal areas.	The activity <b>explicitly</b> calls for students to <b>describe similarities and differences between historical developments, processes, regions, eras, or other focal areas.</b>	The activity <b>explicitly</b> calls for students to describe similarities and differences between historical developments, processes, regions, eras, or other focal areas <b>and prompts students to provide brief explanations of why the similarities and differences existed.</b>	The activity <b>explicitly</b> calls for students to describe similarities and differences between historical developments, processes, regions, eras, or other focal areas and prompts students to provide <b>extended</b> explanations of why the similarities and differences existed.	For example, the activity may prompt students to extend their analysis by <ul style="list-style-type: none"> <li>evaluating the relative historical significance of particular similarities or differences and/or</li> <li>exploring the connection between similarities and differences within and across different categories (e.g., political, religious, geographic).</li> </ul>



**A4. HISTORICAL CONTEXTUALIZATION:** Activity **explicitly calls for students** ~~provides students the opportunity~~ to contextualize historical phenomena and actions within a temporal, spatial and/or sociocultural setting using appropriate historical evidence.

0	1	2	3	Level 3 Examples
The activity <b>does not explicitly</b> call for students to situate phenomena and/or actions in their broader temporal, spatial, and/or sociocultural context.	The activity <b>explicitly</b> calls for students to <b>situate phenomena and/or actions in their broader temporal, spatial, and/or sociocultural context.</b>	The activity <b>explicitly</b> calls for students to situate phenomena and/or actions in their broader temporal, spatial, and/or sociocultural context, <b>and prompts students to provide a brief analysis of how understanding that context improves their ability to interpret the phenomena/action and its significance.</b>	The activity <b>explicitly</b> calls for students to situate phenomena and/or actions in their relevant broader temporal, spatial, and/or sociocultural context and prompts students to provide <b>an extended</b> analysis of how understanding that context improves their ability to interpret the phenomena/action and its significance.	For example, the activity may invite students to acknowledge ways in which contemporary values, attitudes, and conceptualizations differ from those in the past, and show an understanding of how particular perspectives of historical agents would have affected actions.

**A5. CONTINUITY AND CHANGE OVER TIME IN HISTORY**~~HISTORYCOT~~: Activity **explicitly** calls for students~~provides students~~ the opportunity to analyze continuity and change over time using appropriate historical evidence.

0	1	2	3	Level 3 Examples
Activity <b>does not explicitly</b> call for students to analyze continuity and change over time.	Activity <b>explicitly</b> calls for students to <b>provide a description of patterns of continuity and change over time.</b>	The activity <b>explicitly</b> calls for students to describe patterns of continuity and change over time, <b>and asks students to provide a brief analysis of why phenomena persisted or changed.</b>	The activity <b>explicitly</b> calls for students to describe patterns of continuity and change over time, and provide <b>extended</b> analysis of why phenomena persisted or changed.	For example, the activity may prompt students to <ul style="list-style-type: none"> <li>analyze the short-term or long-term historical significance of developments in relation to patterns of change and continuity, and/or</li> <li>ask students to draw conclusions about aspects of patterns, such as their level (global, interregional, regional, or local), speed, and direction (progressive or regressive).</li> </ul>

**A6. SOURCING:** Activity **explicitly calls for students** ~~provides students the opportunity to~~ source a historical document (e.g., identify the author's purpose and perspective).

0	1	2	3	Level 3 Examples
Activity <b>does not explicitly</b> call for students to describe the author's identity and point of view of the author, broader temporal and spatial context, purpose, and/or intended audience of a source.	Activity <b>explicitly</b> calls for students to <b>describe the author's identity and point of view of the author, broader temporal and spatial context, purpose, and/or intended audience of a source.</b>	Activity <b>explicitly</b> calls for students to describe the author's identity and point of view of the author, broader temporal and spatial context, purpose, and/or intended audience of a source <b>and prompts students to briefly analyze how and why some of these factors impacted the way that the author framed the content and how they might affect its meaning.</b>	Activity <b>explicitly</b> calls for students to describe the identity and point of view of the author, broader temporal and spatial context, purpose, and intended audience of a source and prompts <b>an extended analysis of</b> how and why some of these factors impacted the way that the author framed the content and how they might affect its meaning.	For example, the activity may prompt students to evaluate how and why the factors above relate to the historical significance of the source and/or its limitations and credibility.

## Rubric for Examining Opportunities for Historical Thinking Skills in High School World History Student Work

### Introduction

This ~~set of rubrics~~ ~~are~~ ~~meant to be used to study the extent to which high school world history student work exhibits historical thinking skills.~~

**The student work dimensions (S1 - S6) can be used to assess the extent to which student work provides evidence that student:**

- Made, ~~supported or evaluated a historical claim~~ ~~a claim and/or assessed the quality of a claim found in a historical account or interpretation~~ (S1 - ~~Historical argumentation~~ ~~Claims and evidence~~)
- Employed ~~historical~~ causal reasoning (S2 - ~~Historical c~~ ~~G~~ausation)
- Described and explained similarities and differences between historical developments, processes, regions, eras, or other focal areas (S3 - ~~Historical c~~ ~~C~~omparison)
- Contextualized historical phenomena and actions within a temporal, spatial and/or sociocultural setting (S4 - ~~Historical~~ ~~C~~ontextualization)
- Analyzed continuity and change over time ~~in history~~ (S5 - Continuity and ~~c~~ ~~C~~hange ~~o~~ ~~O~~ver ~~t~~ ~~T~~ime ~~in history~~)
- Sourced a historical document (e.g., identified the author's purpose and perspective) (S6 - Sourcing)

Each dimension has four levels (0-3) where generally 0 indicates the absence of a historical thinking practice or skill, 1 indicates emergence, 2 indicates solid presence, and 3 indicates rigorous presence.

**S1. HISTORICAL ARGUMENTATION CLAIMS AND EVIDENCE:** ~~Students work provides evidence that the student made, supported or assessed a historical and developed a claim and/or assessed the quality of a claim found in a historical account or interpretation.~~

0	1	2	3	Level 3 Examples
<p>The student <b>neither makes, supports or assesses a historical claim-a claim nor evaluates the quality and veracity of a claim in a historical account and/or interpretation.</b></p>	<p>The student <b>states a historical claim, a reason for a historical claim or an evaluation of a historical claim makes, a claim and/or evaluates the quality and veracity of a claim in a historical account and/or interpretation.</b></p>	<p>The student <b>states a historical claim, a reason for a historical claim or an evaluation of a historical claim makes a claim and/or evaluates the quality and veracity of a claim in a historical account and/or interpretation.</b></p> <p>The student <b>briefly explains their claim, reason or evaluation. uses evidence and reasoning to support their own claim, and/or to support, qualify, or refute the validity of another person's claim.</b></p>	<p>The student <b>states a historical claim, a reason for a historical claim or an evaluation of a historical claim makes a claim and/or evaluates the quality and veracity of a claim in a historical account and/or interpretation.</b></p> <p>The student <b>provides an extended explanation of their claim, reason or evaluation. makes an extended argument using evidence and reasoning to support their own claim, and/or to support, qualify, or refute the validity of another person's claim.</b></p>	<p>For example, the student may support their own claim by</p> <ul style="list-style-type: none"> <li>• integrating relevant evidence from multiple sources,</li> <li>• acknowledging the credibility and limitations of the evidence used and noting discrepancies across sources,</li> <li>• addressing and evaluating potential counterarguments, and/or</li> <li>• using reasoning to connect the evidence to their claim.</li> </ul> <p>Or, the student may extend their argument over the validity of another person's claim by</p> <ul style="list-style-type: none"> <li>• examining supporting and refuting evidence, including information the claim maker has omitted,</li> <li>• explaining the strengths and limitations of the reasoning used to support the claim,</li> <li>• corroborating the claim with other accounts or interpretations, and/or</li> <li>• considering the perspective and credibility of the claim maker.</li> </ul>

*Note: Possible student misconceptions about claims and evidence include the student grounding their claim in unsupported opinion and/or not distinguishing between historical fact and historical interpretation.*

**S2. HISTORICAL CAUSATION:** ~~Student's work provides evidence that the student~~ employed **historical** causal reasoning using appropriate historical evidence.

0	1	2	3	Level 3 Examples
<p>The student <b>does not</b> describe <b>historical</b> causes and/or effects to explain human actions, events, and/or larger structures or processes.</p> <p><b>OR</b></p> <p>Student employs causal reasoning in non-historical contexts.</p>	<p>The student <b>describes historical causes and/or effects to explain human actions, events, and/or larger structures or processes.</b></p>	<p>The student describes <b>historical</b> causes and/or effects to explain human actions, events, and/or larger structures or processes, <b>and provides a brief analysis of relationships or distinctions between different causes and/or effects.</b></p>	<p>The student describes <b>historical</b> causes and/or effects to explain human actions, events, and/or larger structures or processes, and provides <b>an extended</b> analysis of relationships or distinctions between different causes and/or effects.</p>	<p>For example, the student may</p> <ul style="list-style-type: none"> <li>analyze the distinctions between background conditions, triggering events, primary and secondary causes, and/or immediate and long-term effects.</li> <li>evaluate the relative historical significance of various causes and effects.</li> </ul>

*Notes: Possible student misconceptions about causation include the student conflating causes, actions, and events, believing that a longer list of causes made an event more likely to occur, considering the alternative of a cause to be the lack of an occurrence rather than an alternative occurrence, and/or placing causes in a linear order and arguing that the first cause impacted the second cause and so on, until the event or process occurred.*

**S3. HISTORICAL COMPARISON:** Student ~~work provides evidence that the student~~ described and explained similarities and differences between historical developments, processes, regions, eras, or other focal areas, using appropriate historical evidence.

0	1	2	3	Level 3 Examples
The student <b>does not</b> describe similarities and differences of developments, processes, regions, eras, or other focal areas.	The student <b>describes similarities and differences of developments, processes, regions, eras, or other focal areas.</b>	The student describes similarities and differences between the foci of comparison, <b>and provides a brief analysis of reasons for these similarities and differences.</b>	The student describes similarities and differences between the foci of comparison, and provides <b>an extended</b> analysis of the reasons for these similarities and differences.	For example, the student may <ul style="list-style-type: none"> <li>• extend their analysis by evaluating the relative historical significance of particular similarities or differences and/or</li> <li>• exploring the connection between similarities and differences within and across different categories (e.g., political, religious, geographic).</li> </ul>

**S4. HISTORICAL CONTEXTUALIZATION:** ~~Students work provides evidence that the student~~ contextualized historical phenomena and actions within a temporal, spatial and/or sociocultural setting using appropriate historical evidence.

0	1	2	3	Level 3 Examples
The student <b>does not</b> situate phenomena and/or actions in their broader temporal, spatial, and/or sociocultural context.	The student <b>situates phenomena and/or actions in their broader temporal, spatial, and/or sociocultural context.</b>	The student situates phenomena and/or actions in their broader temporal, spatial, and/or sociocultural context, <b>and provides a brief analysis of how context improves their ability to interpret the phenomena/action and its significance.</b>	The student situates phenomena and/or actions in their relevant broader temporal, spatial, and/or sociocultural context and provides <b>an extended analysis</b> of how understanding that context improves their ability to interpret the phenomena/action and its significance.	For example, the student's connections to context may acknowledge ways in which contemporary values, attitudes, and conceptualizations differ from those in the past, and show an understanding of how particular perspectives of historical agents would have affected actions.

*Note: Possible student misconceptions about contextualization include the student using a present-oriented perspective in thinking about past phenomena and actions.*



**S5. CHANGE AND CONTINUITY OVER TIME IN HISTORY CCOT:** ~~Students work provides evidence that the student analyzed~~  
continuity and change over time using appropriate historical evidence.

0	1	2	3	Level 3 Examples
The student <b>does not</b> describe continuity and change over time.	The student <b>describes patterns of continuity and change over time.</b>	The student describes patterns of continuity and change over time, <b>and provides a brief analysis of why phenomena persisted or changed.</b>	The student describes and explains patterns of continuity and change over time, and provides <b>an extended</b> analysis of why phenomena persisted or changed.	For example, the student may <ul style="list-style-type: none"> <li>analyze the short-term or long-term historical significance of developments and relate them to the larger patterns of change and continuity, and/or</li> <li>draw conclusions about aspects of patterns such as the level (global, interregional, regional, or local), speed, and direction of the change or continuity (progressive or regressive).</li> </ul>

*Note: Possible student misconceptions about change and continuity over time include the student confusing continuity with “no change” occurring, conflating any differences that happened over time as changes, seeing events and changes as synonymous (rather than taking into account gradual change or changes in opinion, circumstance, etc.), conceptualizing all change as progressive, and/or looking at the past through a deficit lens.*

**S6. SOURCING:** Student ~~s work provides evidence that the student~~ sourced a historical document (e.g., identify the author's purpose and perspective).

0	1	2	3	Level 3 Examples
Student <b>does not</b> describe the identity and point of view of the author, broader temporal and spatial context, purpose, and/or intended audience of a source.	Student <b>describes the identity and point of view of the author, broader temporal and spatial context, purpose, and/or intended audience of a source.</b>	Student describes the identity and point of view of the author, broader temporal and spatial context, purpose, and/or intended audience of a source <b>and provides a brief analysis of how some of these factors impacted the way that the author framed the content and how they might affect its meaning.</b>	Student describes the identity and point of view of the author, broader temporal and spatial context, purpose, and intended audience of a source and provides <b>an extended</b> analysis of how and why some of these factors impacted the way that the author framed the content and how they might affect its meaning.	For example, the student may evaluate how and why the factors above relate to the historical significance of the source and/or its limitations and credibility.

**Appendix J. Final rubrics**

## **Rubrics for Examining Opportunities for Historical Thinking Practices in High School World History Activities**

### **Introduction**

This set of rubrics are meant to be used to study the extent to which high school world history classroom activities provide opportunities for students to engage in historical thinking practices.

**The activity dimensions A1 - A6 can be used to assess the extent to which a learning activity provides students the opportunity to:**

- Make and develop claims and/or assess the quality of claims found in a historical account or interpretation (A1 - Historical argumentation)
- Employ causal reasoning (A2 - Historical causation)
- Describe and explain similarities and differences between historical developments, processes, regions, eras, or other focal areas (A3 - Historical comparison)
- Contextualize historical phenomena and actions within a temporal, spatial and/or sociocultural setting (A4 - Historical contextualization)
- Analyze continuity and change over time in history (A5 - Continuity and change over time in History)
- Source a historical document (e.g., identify the author's purpose and perspective) (A6 - Sourcing)

Each dimension has four levels (0-3) where generally 0 indicates the absence of a historical thinking practice in the activity, 1 indicates emergence, 2 indicates partial presence, and 3 indicates solid presence.

**A1. HISTORICAL ARGUMENTATION: Activity explicitly calls for students to make, support or assess a historical claim.**

0	1	2	3	Level 3 Examples
Activity <b>does not explicitly</b> call for students to make, support or assess a historical claim.	The activity <b>explicitly prompts students to state</b> a historical claim, a reason for a historical claim or an evaluation of a historical claim .	<p>The activity explicitly prompts students to state a historical claim, a reason for a historical claim or an evaluation of a historical claim .</p> <p><b>The activity explicitly prompts students to briefly explain their claim, reason, or evaluation.</b></p>	<p>The activity explicitly prompts students to state a historical claim, a reason for a historical claim or an evaluation of a historical claim.</p> <p>The activity explicitly prompts students to <b>provide an extended explanation</b> of their claim, reason, or evaluation.</p>	<p>For example, the activity may prompt students to support their own claim by</p> <ul style="list-style-type: none"> <li>• integrating relevant evidence from multiple sources,</li> <li>• acknowledging the credibility and limitations of the evidence used and noting discrepancies across sources,</li> <li>• addressing and evaluating potential counterarguments, and/or</li> <li>• using reasoning to connect the evidence to their claim.</li> </ul> <p>Or, the activity may prompt students to extend their argument on the validity of another person's claim by</p> <ul style="list-style-type: none"> <li>• examining supporting and refuting evidence, including information the claim maker has omitted,</li> <li>• explaining the strengths and limitations of the reasoning used to support the claim,</li> <li>• corroborating the claim with other accounts or interpretations, and/or</li> <li>• considering the perspective and credibility of the claim maker.</li> </ul>

\*Historical claims can include historical accounts and interpretations.

**A2. HISTORICAL CAUSATION: Activity explicitly calls for students to employ causal reasoning using appropriate historical evidence.**

0	1	2	3	Level 3 Examples
The activity <b>does not explicitly</b> call for students to describe causes and/or effects to explain human actions, events, and/or larger structures or processes.	The activity explicitly calls for students to <b>describe causes and/or effects to explain human actions, events, and/or larger structures or processes.</b>	The activity explicitly calls for students to describe causes and/or effects to explain human actions, events, and/or larger structures or processes, <b>and asks students to provide a brief analysis of distinctions between different causes and/or effects (e.g., primary vs. secondary or immediate vs. long-term) or the relationship between causes and/or effects.</b>	The activity explicitly calls for students to describe causes and/or effects to explain human actions, events, and/or larger structures or processes, and asks students to provide <b>an extended</b> analysis of distinctions between different causes and/or effects (e.g., primary vs. secondary or immediate vs. long-term) or the relationship between causes and/or effects.	For example, the activity may prompt students to <ul style="list-style-type: none"> <li>• provide an extended analysis of the distinctions between background conditions, triggering events, primary and secondary causes, and/or immediate and long-term effects.</li> <li>• evaluate the relative historical significance of various causes and effects.</li> </ul>

**A3. HISTORICAL COMPARISON:** Activity explicitly calls for students to describe and explain similarities and differences between historical developments and processes, regions, eras, or other focal areas, using appropriate historical evidence.

0	1	2	3	Level 3 Examples
Activity <b>does not explicitly</b> call for students to describe similarities and differences between historical developments, processes, regions, eras, or other focal areas.	The activity explicitly calls for students to <b>describe similarities and differences between historical developments, processes, regions, eras, or other focal areas.</b>	The activity explicitly calls for students to describe similarities and differences between historical developments, processes, regions, eras, or other focal areas <b>and prompts students to provide brief explanations of why the similarities and differences existed.</b>	The activity explicitly calls for students to describe similarities and differences between historical developments, processes, regions, eras, or other focal areas and prompts students to provide <b>extended</b> explanations of why the similarities and differences existed.	For example, the activity may prompt students to extend their analysis by <ul style="list-style-type: none"> <li>evaluating the relative historical significance of particular similarities or differences and/or</li> <li>exploring the connection between similarities and differences within and across different categories (e.g., political, religious, geographic).</li> </ul>

**A4. HISTORICAL CONTEXTUALIZATION:** Activity explicitly calls for students to contextualize historical phenomena and actions within a temporal, spatial and/or sociocultural setting using appropriate historical evidence.

0	1	2	3	Level 3 Examples
The activity <b>does not</b> explicitly call for students to situate phenomena and/or actions in their broader temporal, spatial, and/or sociocultural context.	The activity explicitly calls for students to <b>situate phenomena and/or actions in their broader temporal, spatial, and/or sociocultural context.</b>	The activity explicitly calls for students to situate phenomena and/or actions in their broader temporal, spatial, and/or sociocultural context, <b>and prompts students to provide a brief analysis of how understanding that context improves their ability to interpret the phenomena/action and its significance.</b>	The activity explicitly calls for students to situate phenomena and/or actions in their relevant broader temporal, spatial, and/or sociocultural context and prompts students to provide <b>an extended</b> analysis of how understanding that context improves their ability to interpret the phenomena/action and its significance.	For example, the activity may invite students to acknowledge ways in which contemporary values, attitudes, and conceptualizations differ from those in the past, and show an understanding of how particular perspectives of historical agents would have affected actions.



**A5. CONTINUITY AND CHANGE OVER TIME IN HISTORY:** Activity explicitly calls for students to analyze continuity and change over time using appropriate historical evidence.

0	1	2	3	Level 3 Examples
Activity <b>does not</b> explicitly call for students to analyze continuity and change over time.	Activity explicitly calls for students to <b>provide a description of patterns of continuity and change over time.</b>	The activity explicitly calls for students to describe patterns of continuity and change over time, <b>and asks students to provide a brief analysis of why phenomena persisted or changed.</b>	The activity explicitly calls for students to describe patterns of continuity and change over time, and provide <b>extended</b> analysis of why phenomena persisted or changed.	For example, the activity may prompt students to <ul style="list-style-type: none"> <li>analyze the short-term or long-term historical significance of developments in relation to patterns of change and continuity, and/or</li> <li>ask students to draw conclusions about aspects of patterns, such as their level (global, interregional, regional, or local), speed, and direction (progressive or regressive).</li> </ul>

**A6. SOURCING:** Activity explicitly calls for students source a historical document (e.g., identify the author’s purpose and perspective).

0	1	2	3	Level 3 Examples
Activity <b>does not</b> explicitly call for students to describe the author’s identity and point of view of the author, broader temporal and spatial context, purpose, and/or intended audience of a source.	Activity explicitly calls for students to <b>describe the author’s identity and point of view of the author, broader temporal and spatial context, purpose, and/or intended audience of a source.</b>	Activity explicitly calls for students to describe the author’s identity and point of view of the author, broader temporal and spatial context, purpose, and/or intended audience of a source <b>and prompts students to briefly analyze how and why some of these factors impacted the way that the author framed the content and how they might affect its meaning.</b>	Activity explicitly calls for students to describe the identity and point of view of the author, broader temporal and spatial context, purpose, and intended audience of a source and prompts <b>an extended analysis of</b> how and why some of these factors impacted the way that the author framed the content and how they might affect its meaning.	For example, the activity may prompt students to evaluate how and why the factors above relate to the historical significance of the source and/or its limitations and credibility.

## **Rubric for Examining Opportunities for Historical Thinking Skills in High School World History Student Work**

### **Introduction**

This set of rubrics are meant to be used to study the extent to which high school world history student work exhibits historical thinking skills.

**The student work dimensions (S1 - S6) can be used to assess the extent to which student work provides evidence that student:**

- Made, supported or evaluated a historical claim (S1 - Historical argumentation)
- Employed historical causal reasoning (S2 - Historical causation)
- Described and explained similarities and differences between historical developments, processes, regions, eras, or other focal areas (S3 - Historical comparison)
- Contextualized historical phenomena and actions within a temporal, spatial and/or sociocultural setting (S4 - Historical contextualization)
- Analyzed continuity and change over time in history (S5 - Continuity and change over time in history)
- Sourced a historical document (e.g., identified the author's purpose and perspective) (S6 - Sourcing)

Each dimension has four levels (0-3) where generally 0 indicates the absence of a historical thinking practice or skill, 1 indicates emergence, 2 indicates solid presence, and 3 indicates rigorous presence.

**S1. HISTORICAL ARGUMENTATION: Students made, supported or assessed a historical claim.**

0	1	2	3	Level 3 Examples
The student <b>neither makes, supports or assesses a historical claim.</b>	The student <b>states a historical claim, a reason for a historical claim or an evaluation of a historical claim.</b>	<p>The student states a historical claim, a reason for a historical claim or an evaluation of a historical claim.</p> <p>The student <b>briefly explains their claim, reason or evaluation.</b></p>	<p>The student states a historical claim, a reason for a historical claim or an evaluation of a historical claim.</p> <p>The student provides <b>an extended explanation</b> of their claim, reason or evaluation.</p>	<p>For example, the student may support their own claim by</p> <ul style="list-style-type: none"> <li>• integrating relevant evidence from multiple sources,</li> <li>• acknowledging the credibility and limitations of the evidence used and noting discrepancies across sources,</li> <li>• addressing and evaluating potential counterarguments, and/or</li> <li>• using reasoning to connect the evidence to their claim.</li> </ul> <p>Or, the student may extend their argument over the validity of another person's claim by</p> <ul style="list-style-type: none"> <li>• examining supporting and refuting evidence, including information the claim maker has omitted,</li> <li>• explaining the strengths and limitations of the reasoning used to support the claim,</li> <li>• corroborating the claim with other accounts or interpretations, and/or</li> <li>• considering the perspective and credibility of the claim maker.</li> </ul>

*Note: Possible student misconceptions about claims and evidence include the student grounding their claim in unsupported opinion and/or not distinguishing between historical fact and historical interpretation.*

**S2. HISTORICAL CAUSATION: Students employed historical causal reasoning using appropriate historical evidence.**

0	1	2	3	Level 3 Examples
<p>The student <b>does not</b> describe historical causes and/or effects to explain human actions, events, and/or larger structures or processes.</p> <p>OR</p> <p>Student employs causal reasoning in non-historical contexts.</p>	<p>The student <b>describes historical causes and/or effects to explain human actions, events, and/or larger structures or processes.</b></p>	<p>The student describes historical causes and/or effects to explain human actions, events, and/or larger structures or processes, <b>and provides a brief analysis of relationships or distinctions between different causes and/or effects.</b></p>	<p>The student describes historical causes and/or effects to explain human actions, events, and/or larger structures or processes, and provides <b>an extended</b> analysis of relationships or distinctions between different causes and/or effects.</p>	<p>For example, the student may</p> <ul style="list-style-type: none"> <li>analyze the distinctions between background conditions, triggering events, primary and secondary causes, and/or immediate and long-term effects.</li> <li>evaluate the relative historical significance of various causes and effects.</li> </ul>

*Notes: Possible student misconceptions about causation include the student conflating causes, actions, and events, believing that a longer list of causes made an event more likely to occur, considering the alternative of a cause to be the lack of an occurrence rather than an alternative occurrence, and/or placing causes in a linear order and arguing that the first cause impacted the second cause and so on, until the event or process occurred.*

**S3. HISTORICAL COMPARISON:** Students described and explained similarities and differences between historical developments, processes, regions, eras, or other focal areas, using appropriate historical evidence.

0	1	2	3	Level 3 Examples
The student <b>does not</b> describe similarities and differences of developments, processes, regions, eras, or other focal areas.	The student <b>describes similarities and differences of developments, processes, regions, eras, or other focal areas.</b>	The student describes similarities and differences between the foci of comparison, <b>and provides a brief analysis of reasons for these similarities and differences.</b>	The student describes similarities and differences between the foci of comparison, and provides <b>an extended</b> analysis of the reasons for these similarities and differences.	For example, the student may <ul style="list-style-type: none"> <li>• extend their analysis by evaluating the relative historical significance of particular similarities or differences and/or</li> <li>• exploring the connection between similarities and differences within and across different categories (e.g., political, religious, geographic).</li> </ul>

**S4. HISTORICAL CONTEXTUALIZATION:** Students contextualized historical phenomena and actions within a temporal, spatial and/or sociocultural setting using appropriate historical evidence.

0	1	2	3	Level 3 Examples
The student <b>does not</b> situate phenomena and/or actions in their broader temporal, spatial, and/or sociocultural context.	The student <b>situates phenomena and/or actions in their broader temporal, spatial, and/or sociocultural context.</b>	The student situates phenomena and/or actions in their broader temporal, spatial, and/or sociocultural context, <b>and provides a brief analysis of how context improves their ability to interpret the phenomena/action and its significance.</b>	The student situates phenomena and/or actions in their relevant broader temporal, spatial, and/or sociocultural context and provides <b>an extended analysis</b> of how understanding that context improves their ability to interpret the phenomena/action and its significance.	For example, the student's connections to context may acknowledge ways in which contemporary values, attitudes, and conceptualizations differ from those in the past, and show an understanding of how particular perspectives of historical agents would have affected actions.

*Note: Possible student misconceptions about contextualization include the student using a present-oriented perspective in thinking about past phenomena and actions.*

**S5. CHANGE AND CONTINUITY OVER TIME IN HISTORY: Students analyzed continuity and change over time using appropriate historical evidence.**

0	1	2	3	Level 3 Examples
The student <b>does not</b> describe continuity and change over time.	The student <b>describes patterns of continuity and change over time.</b>	The student describes patterns of continuity and change over time, <b>and provides a brief analysis of why phenomena persisted or changed.</b>	The student describes and explains patterns of continuity and change over time, and provides <b>an extended</b> analysis of why phenomena persisted or changed.	<p>For example, the student may</p> <ul style="list-style-type: none"> <li>analyze the short-term or long-term historical significance of developments and relate them to the larger patterns of change and continuity, and/or</li> <li>draw conclusions about aspects of patterns such as the level (global, interregional, regional, or local), speed, and direction of the change or continuity (progressive or regressive).</li> </ul>

*Note: Possible student misconceptions about change and continuity over time include the student confusing continuity with “no change” occurring, conflating any differences that happened over time as changes, seeing events and changes as synonymous (rather than taking into account gradual change or changes in opinion, circumstance, etc.), conceptualizing all change as progressive, and/or looking at the past through a deficit lens.*



**S6. SOURCING: Students source a historical document (e.g., identify the author's purpose and perspective).**

0	1	2	3	Level 3 Examples
Student <b>does not</b> describe the identity and point of view of the author, broader temporal and spatial context, purpose, and/or intended audience of a source.	Student <b>describes the identity and point of view of the author, broader temporal and spatial context, purpose, and/or intended audience of a source.</b>	Student describes the identity and point of view of the author, broader temporal and spatial context, purpose, and/or intended audience of a source <b>and provides a brief analysis of how some of these factors impacted the way that the author framed the content and how they might affect its meaning.</b>	Student describes the identity and point of view of the author, broader temporal and spatial context, purpose, and intended audience of a source and provides <b>an extended</b> analysis of how and why some of these factors impacted the way that the author framed the content and how they might affect its meaning.	For example, the student may evaluate how and why the factors above relate to the historical significance of the source and/or its limitations and credibility.