

What Is Learned in College History Classes?

Sam Wineburg, Mark Smith, and Joel Breakstone

Historians are great at telling stories to others. But they also tell stories to themselves, including one that says they “suck at assessment”; an essay by Anne Hyde, in the “Textbooks and Teaching” section of a previous *Journal of American History* issue, proclaimed as much. Hyde, the 2012 Bancroft Prize winner for *Empires, Nations, and Families*, told a story about what happened when Colorado College failed its accreditation review. Until that unexpected blow, Hyde’s colleagues had dismissed accreditors’ requests for evidence of learning in the major. Beyond reporting department enrollments and student grade point averages, the department acted as independent contractors: faculty members were deeply committed to their own particular courses and treated their classrooms, as Hyde put it, “as private, sacred spaces.” Faculty could go on at length about their assigned readings and course expectations, but collectively they “had no clue about how it all added up” to form something greater than an aggregation of disparate puzzle pieces.¹

This situation is not unique to Colorado College. The American Historical Association (AHA) has recognized that many departments were wrestling with similar issues. In response, the AHA initiated the History Tuning Project, which sought to “describe the skills, knowledge, and habits of mind that students develop in history courses and degree programs.” The project’s goal is to foster collaboration among college faculty by providing a framework to “lay out their own distinctive goals and outcomes.” Although the AHA Tuning Project made progress in identifying history’s core concepts, the AHA laid out challenges that remain for the field in its history discipline core analysis for 2016: “How do we know our students are learning the outcomes laid out here? What are

Sam Wineburg is the Margaret Jacks Professor of Education and of History at Stanford University. Mark Smith is the director of assessment at the Stanford History Education Group. Joel Breakstone is the director of the Stanford History Education Group.

The Library of Congress’s Teaching with Primary Sources Program supported the Stanford History Education Group’s development of new history assessments and the creation of a Web site to host them, *Beyond the Bubble*, <https://sheg.stanford.edu/history-assessments>. Content created and featured in partnership with the Teaching with Primary Sources Program does not indicate an endorsement by the Library of Congress.

Readers may contact Wineburg at wineburg@stanford.edu. Readers may contact Smith at msmith4@stanford.edu. Readers may contact Breakstone at breakstone@stanford.edu.

¹ Anne Hyde, “Five Reasons History Professors Suck at Assessment,” *Journal of American History*, 102 (March 2016), 1104–7, esp. 1105.

the meaningful ways we can demonstrate that students have in fact achieved the expectations we set for them?”²

Our work has addressed these questions. Over the past seven years, we have engaged in research and development to create assessments that measure historical thinking. Although our work has mostly focused on high schools, here we present results from a study of what happened when we gave our assessments to college students, majors and nonmajors alike.

History Assessments of Thinking

The challenge of how to measure learning is not restricted to universities. For high school teachers the situation is not much better. The structure of the school day restricts collaboration to brief meetings taken up by administrative matters, leaving scant time for teachers to articulate goals for student learning. Moreover, few options exist for assessing student learning. Multiple-choice tests dominate at the high school level. Each of the twenty-four states that test students in history uses multiple-choice questions and over half use only multiple-choice questions. Analytic essays rank a close second to multiple-choice questions as testing options. These essays provide students opportunities to practice skills central to the discipline, but as assessment tools they are blunt instruments: so many processes occur at once that it is hard to know what, exactly, these tasks measure. From the perspective of cognitive science, pinpointing the factors that go into an essay of the sort used in the College Board's Advanced Placement program's "document-based question" (DBQ) is virtually impossible. Even after decades of developing and refining the DBQ, reliability (that is, the degree of consistency in test scores) remains disturbingly low.³

With support from the Library of Congress, we developed dozens of tasks for assessing historical thinking at the high school level. Our tasks ask students to answer questions about historical sources and to explain their reasoning in a few sentences. Each task assesses one or more historical thinking "constructs"—core notions of historical thinking, such as the relationship between claim and evidence, the nature of chronological thinking, or how time and place influence events. These aspects apply whether one is reasoning about why Constantine converted to Christianity in 312 or why World War I erupted in 1914.⁴

For example, one of our tasks presents students with excerpts from two documents about the Philippine-American War and asks how each provides evidence of opposition to the war. One source is sworn testimony by the U.S. Army corporal Richard O'Brien before the Senate Committee on the Philippines, chaired in 1902 by the Massachusetts

² "AHA History Tuning Project: 2016 Discipline Core," Dec. 2016, *American Historical Association*, <https://www.historians.org/teaching-and-learning/tuning-the-history-discipline/2016-history-discipline-core>.

³ Pam Grossman, Sam Wineburg, and Stephen Woolworth, "Toward a Theory of Teacher Community," *Teachers College Record*, 103 (Dec. 2001), 942–1012. Daisy Martin et al., *A Report on the State of History Education: State Policies and National Programs* (Fairfax, 2011); Sam Wineburg, "Crazy for History," *Journal of American History*, 90 (March 2004), 1401–14. Howard Wainer, *Uneducated Guesses: Using Evidence to Uncover Misguided Education Policies* (Princeton, 2011), 109.

⁴ Joel Breakstone, "Try, Try, Try Again: The Process of Designing New History Assessments," *Theory & Research in Social Education*, 42 (no. 4, 2014), 453–85; Joel Breakstone, Mark Smith, and Sam Wineburg, "Beyond the Bubble in History/Social Studies Assessments," *Phi Delta Kappan*, 94 (Feb. 2013), 53–57; Mark Smith, "Cognitive Validity: Can Multiple-Choice Questions Tap Historical Thinking Processes?," *American Educational Research Journal*, 54 (Dec. 2017), 1011–47; Sam Wineburg, Mark Smith, and Joel Breakstone, "New Directions in Assessment: Using Library of Congress Sources to Assess Historical Understanding," *Social Education*, 76 (Nov.–Dec. 2012), 290–93; Sam Wineburg, *Why Learn History When It's Already on Your Phone?* (Chicago, 2018).

Figure 1
Opposition to Philippine-American War Assessment

How does each document provide evidence that many Americans opposed the war?

Document A: *The following is an excerpt from sworn testimony given before the U.S. Senate by Corporal Richard O'Brien in 1902. O'Brien was called to testify in a Senate investigation of alleged war crimes committed by American soldiers in the Philippine-American War.*

"We entered the town. It was just daybreak. The first thing we saw was a boy coming down . . . and the first sergeant, William Stahlburg, shot at the boy . . . That brought the people in the houses out, brought them to the doors and out into the street, and how the order started and who gave it I don't know, but the town was fired on . . . After that two old men came out, hand in hand. I should think they were over 50 years old, probably between 50 and 70 years old. They had a white flag. They were shot down. At the other end of the town we heard screams, and there was a woman there; she was burned up, and in her arms was a baby . . . There was not a shot fired on the part of the Filipinos."

Document B: *The following is an excerpt from a letter by Colonel Frederick Funston that was published in the Kansas City Journal on April 22, 1899. Funston, who was a war hero for his extensive service in the Philippine-American War, wrote and spoke often about the Philippine-American War in order to increase public support for American involvement in the conflict.*

"I am afraid that some people at home will lie awake nights worrying about the ethics of this war, thinking that our enemy is fighting for the right of self-government . . . [The Filipinos] have a certain number of educated leaders—educated, however, about the same way a parrot is. They are, as a rule, an illiterate, semi-savage people who are waging war not against tyranny, but against Anglo-Saxon order and decency . . . I, for one, hope that Uncle Sam will apply the chastening rod good, hard and plenty, and lay it on until they come in to the reservation and promise to be good 'Injuns.'"

Shown here is an example of an assessment task asking students to examine the content of two documents while considering the occasions that prompted their creation. SOURCES: For document A, Testimony of Richard T. O'Brien, U.S. Congress, Senate, Committee on the Philippines, *Affairs in the Philippines: Hearings before the Committee on the Philippines of the United States Senate*, 57 Cong., 1 sess., April 2, 1902, pp. 2549–51. For document B, "Interesting Letter from Funston," *Kansas City Journal*, April 22, 1899, *Library of Congress*, <http://chroniclingamerica.loc.gov/lccn/sn86063615/1899-04-22/ed-1/seq-4/>.

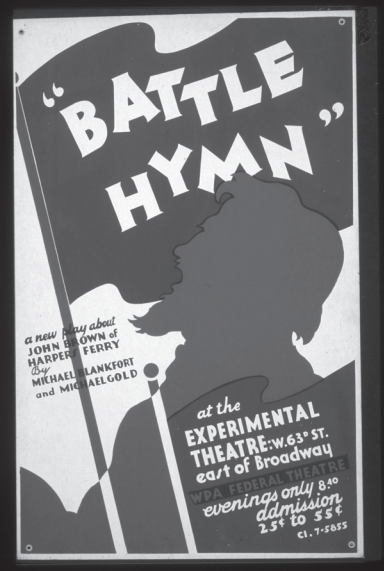
Republican senator Henry Cabot Lodge. The other is from an 1899 letter published in the *Kansas City Journal* by Col. Frederick Funston, who defended American involvement by casting the Filipinos as "illiterate, semi-savage people" who wage war "against Anglo-Saxon order." To succeed in the task, students needed to look beyond the content of the documents to consider the occasions that prompted their creation. Senate committees are not haphazardly convened. High-ranking officers do not write letters defending military campaigns without cause. At its most basic level, this task is about warrant. Students are provided with a claim and evidence, and must specify the relationship between the two.⁵ (See figure 1.)

College Assessment

Our initial work with high school teachers showed promise. Teachers were able to use assessments to gauge students' grasp of key concepts and to inform department-wide discussions about instruction. We began to wonder whether our tasks might address the

⁵ Testimony of Richard T. O'Brien, U.S. Congress, Senate, Committee on the Philippines, *Affairs in the Philippines: Hearings before the Committee on the Philippines of the United States Senate*, 57 Cong., 1 sess., April 2, 1902, pp. 2549–51; "Interesting Letter from Funston," *Kansas City Journal*, April 22, 1899, *Library of Congress*, <http://chroniclingamerica.loc.gov/lccn/sn86063615/1899-04-22/ed-1/seq-4/>.

Figure 2
John Brown Playbill Assessment

<p>Background Information: This is a poster for a play written in 1936 that celebrates the abolitionist John Brown, who tried to start a slave revolt in Harpers Ferry, Virginia, in 1859.</p> 	<p>Question 1: When was the play written?</p> <p>Question 2: Three facts are listed below. Explain whether each fact <u>does or does not</u> provide evidence for why the authors wrote the play.</p> <hr/> <p>Fact 1: Slaves made up nearly 40% of Virginia's population in 1859.</p> <p>Fact 1 does/does not (circle one) provide evidence for why for authors wrote the play because</p> <hr/> <p>Fact 2: One of the play's authors, Michael Gold, was a member of the Communist party, which protested against lynching in the 1930s.</p> <p>Fact 2 does/does not (circle one) provide evidence for why the authors wrote the play because</p> <hr/> <p>Fact 3: After seceding from the Union in 1861, Virginia became the largest state in the Confederacy and the home of its capital, Richmond.</p> <p>Fact 3 does/does not (circle one) provide evidence for why the authors wrote the play because</p> <hr/>
--	---

The assessment task shown here asked students to analyze a document as a historically conscious product of its time. SOURCE: George Goldschmidt, “‘Battle Hymn’ a New Play about John Brown of Harpers Ferry by Michael Blankfort and Michael Gold at the Experimental Theatre,” ca. 1936–1941, *Library of Congress Prints and Photographs Online Catalog*, <http://www.loc.gov/pictures/item/98516478/>.

AHA History Tuning Project’s call for measures to assess the discipline’s core concepts at the college level. Of the thousands of high school students who completed our assessments, most struggled. Would college students exposed to more sophisticated content and a greater range of sources do better? To answer these questions, we administered our tasks to students enrolled in a required introductory U.S. history course at a state university on the West Coast.⁶

In addition to the Philippine-American War task, we gave students a 1936 playbill for *Battle Hymn*, a stage production celebrating John Brown’s 1859 raid on an arsenal at Harpers Ferry, Virginia. Students had to determine whether three facts, each true, might provide evidence for why the authors wrote the play. (See figure 2.) Just as Arthur Miller’s *The Crucible*, about seventeenth-century Salem, Massachusetts, witch trials, reflected the McCarthyism of the 1950s, our task asked students how a play about events in the 1850s might reflect the 1930s. Students struggled with the task in early piloting, but we could not tell if it was because they overlooked the play’s date or thought that the date was ir-

⁶ Joel Breakstone, “History Assessments of Thinking: Design, Interpretation, and Implementation” (Ph.D. diss., Stanford University, 2013), <https://purl.stanford.edu/nt301xp3169>; Joel Breakstone, Sam Wineburg, and Mark Smith, “Formative Assessment Using Library of Congress Documents,” *Social Education*, 79 (Sept. 2015), 178–82; Joel Breakstone and Sam Wineburg, “Ask a Colleague: Formative Assessment,” *ibid.*, 80 (Jan.–Feb. 2016), 8–11.

Figure 3
Thanksgiving Assessment

Directions: Use the painting to answer the question below.



Title: *The First Thanksgiving 1621*

By: J. L. G. Ferris

Date: 1932

STATEMENT: The painting *The First Thanksgiving 1621* helps historians understand the relationship between the Wampanoag Indians and the Pilgrim settlers in 1621.

QUESTION: Do you agree or disagree? (Circle one).

Briefly support your answer:

The assessment task shown here measured students' ability to evaluate a source's evidentiary value from its bibliographic information. SOURCE: Jean Leon Gerome Ferris, *The First Thanksgiving 1621*, ca. 1912–1930, *Library of Congress Prints and Photographs Online Catalog*, <http://www.loc.gov/pictures/item/2001699850/>.

relevant to understanding the authors' motivations. We thus added the first question to make the playbill's date impossible to miss. In subsequent administrations of the task, no student got the date wrong, but most continued to struggle when analyzing the document as a product of its time.⁷

Our third task focused on sourcing: Would students attend to a document's bibliographic information when judging its evidentiary value? We used an early twentieth-century painting, *The First Thanksgiving 1621*, by Jean Leon Gerome Ferris to ask students if the work would be a useful source for historians who wanted to understand the relationship between the Wampanoag and Pilgrim settlers in 1621. (See figure 3.)⁸

⁷ George Goldschmidt, "Battle Hymn' a New Play about John Brown of Harpers Ferry by Michael Blankfort and Michael Gold at the Experimental Theatre," ca. 1936–1941, illustration, *Library of Congress Prints and Photographs Online Catalog*, <http://www.loc.gov/pictures/item/98516478/>. Arthur Miller, *The Crucible: A Play in Four Acts* (New York, 1953).

⁸ Jean Leon Gerome Ferris, *The First Thanksgiving 1621*, ca. 1912–1930, painting, *Library of Congress Prints and Photographs Online Catalog*, <http://www.loc.gov/pictures/item/2001699850/>. The copy of the painting in the exercise was published in 1932.

Study 1

We administered the three exercises at midsemester to seventy-eight freshmen and sophomores in a required U.S. history course. We used a three-point rubric to score responses: “Basic” (zero points) if the answer was off base and bore no relation to the competency being measured; “Emergent” (one point) if the answer showed inklings of proficiency; and “Proficient” (two points) if the answer demonstrated understanding. Across the three assessments were six questions (one for the Ferris painting, two on the Philippine-American War documents, and three on the *Battle Hymn* playbill) resulting in a possible total score of twelve points.

Results were alarming. Students averaged less than one-half of one point. The high mark across the entire sample was a mere three points (earned by three of seventy-eight students). On the painting evaluation task, the average score for students hovered slightly above zero. (See figure 4.) Among students assigned the task, 94 percent ignored the bibliographic information accompanying the picture and evaluated the painting based on whether it matched their preconceptions about Thanksgiving. As one student wrote, “I agree [it would help historians]. The painting does show the nature of the relationship. In the image, we see Pilgrims and Indians interacting peacefully and joyfully.” Other students engaged in a similar matching process but reached the opposite conclusion, rejecting the painting because it conflicted with their prior understanding. As one student explained, “The painting shows a pretty picture of how the Wampanoag Indians and the Pilgrims were sharing a meal and getting along, when in reality the Pilgrims didn’t come and have a peaceful communication. In reality, they came hungry for land and killed or fought anything and anyone trying to stop them.” In neither case did the temporal gap between the image and the event it purports to depict enter into students’ deliberations.

Only one student focused on this gap and provided a rationale for why it mattered: “It was painted in 1932 and the event occurred over 300 years ago. We don’t know if the painter used a credible source to paint the painting and we don’t know if the event even looked like that back then. It’s all speculation from the painter.” This type of reasoning—which we would hope college students would learn to do in an introductory course—was rare.

Based on our experience with high school students, we suspected some college students might struggle. But we woefully underestimated how much they would struggle. Our findings raised questions about the transition from high school to college and the capabilities we can assume that students bring to introductory classes. But what about students in upper-level history courses? Would they breeze through tasks designed for high school students?

Study 2

We administered the same three tasks to forty-nine juniors and seniors enrolled in upper-level history courses at a different state university with a similar student population. Each student had completed at least five university history courses, and twenty-seven of the forty-nine were history majors.

Figure 4
Sample Student Response to the Thanksgiving Task

STATEMENT: The painting *The First Thanksgiving 1621* helps historians understand the relationship between the Wampanoag Indians and the Pilgrim settlers in 1621.

QUESTION: Do you agree or disagree? (Circle one).

Briefly support your answer:

I agree because in the picture it shows the wampanoag Indians and the pilgrims interacting. It shows the Pilgrims sharing their food with the Indians and The Kindness of the pilgrim people.

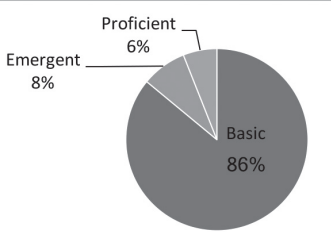
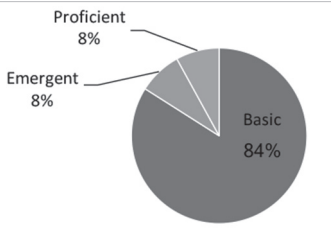
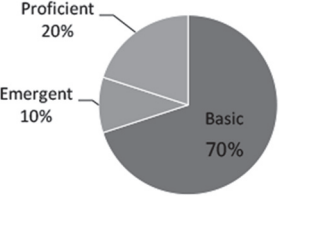
This handwritten paragraph is a typical student response to the task of evaluating a source's evidentiary value from its bibliographic information. Although many aspects of the painting are worthy of analysis, our task focused on one aspect of sourcing: the date. We registered a variety of responses that represented progress toward proficiency. For example, if students did not mention the temporal gap but speculated on the motivations of the artist, we granted partial credit. Responses awarded full credit had to note the gap in time between the creation of the painting and the event it depicts.

Recall that the Philippine-American War task asked students to explain how testimony from a Senate hearing and a letter from a U.S. Army colonel provided evidence of opposition to the war. If students explained in basic contour how each of the two documents provided evidence of public opposition, they earned a total of four points.

These juniors and seniors scored, on average, less than one of four possible points (.77). Eighty-six percent earned no credit on the question about the Senate testimony. Rather than consider what prompted a congressional investigation, students fixated on the atrocities described by Corporal O'Brien in his statement. One history major wrote, "Well, provided what occurred in Document A is true, then it makes sense Americans would oppose the war. Document A would be something someone would quote who opposed the war." Another wrote: "It appears that the lower end of the chain of command was against the war in the Philippines. Due to brutal means of handling the situation in the Philippines many Americans were appalled by such actions." Another wrote, "Many Americans would oppose a war in which the opposing forces did not shoot a single bullet and came out waving a white flag. Americans generally have a difficult time dealing with the murder of children." Students ignored the context of the testimony and focused solely on its content. Of these forty-nine juniors and seniors, only three provided explanations that considered the context of the testimony. One of them wrote, "[The testimony] provides evidence that many Americans opposed the war by there being a Senate investigation. If there hadn't been such a huge opposition by Americans to this war, I don't believe that the investigation would have occurred."

Students did only slightly better on the second question. Over four-fifths failed to note that Colonel Funston was likely responding to public opposition or that the letter's appearance in a newspaper signaled a broader debate about the war. For some, Funston's letter provided no evidence of public opposition. One student reasoned that the letter "does not provide evidence that many Americans opposed the war . . . it's an opinion of a man who supported the war." Other students could not get past Funston's racism. One major

Figure 5
Sample Responses from Students in Upper-Level History Courses

Task	Basic	Emergent	Proficient								
<p><i>Philippine-American War</i> (Question 1)</p> <p>How does the document provide evidence that many Americans opposed the war?</p>  <table><tr><th>Level</th><th>Percentage</th></tr><tr><td>Basic</td><td>86%</td></tr><tr><td>Emergent</td><td>8%</td></tr><tr><td>Proficient</td><td>6%</td></tr></table>	Level	Percentage	Basic	86%	Emergent	8%	Proficient	6%	<p>“The descriptions are of an atrocity, in which Americans killed innocents. This is a very negative image being used to get people to further oppose the use of force in the Philippines.”</p>	<p>“Well it doesn’t, but I guess the fact that there is a U.S. soldier testifying at Senate hearing for alleged war crimes means that some people disliked the war. The corporal that testified clearly felt that he saw war crimes (murdering unarmed civilians), however nowhere does it say he opposed the war.”</p>	<p>“Document A provides evidence that many Americans opposed the war by there being a Senate investigation. If there hadn’t been such a huge opposition by Americans of this war, I don’t believe that the investigation would have occurred.”</p>
Level	Percentage										
Basic	86%										
Emergent	8%										
Proficient	6%										
<p><i>John Brown Playbill</i></p> <p>Fact 1: Slaves made up nearly 40% of Virginia’s population in 1859.</p> <p>Explain whether the fact <u>does</u> or <u>does not</u> provide evidence for why the authors wrote the play.</p>  <table><tr><th>Level</th><th>Percentage</th></tr><tr><td>Basic</td><td>84%</td></tr><tr><td>Emergent</td><td>8%</td></tr><tr><td>Proficient</td><td>8%</td></tr></table>	Level	Percentage	Basic	84%	Emergent	8%	Proficient	8%	<p>“Fact 1 does provide evidence for why the authors wrote the play because it shows that it’s an issue in the country that needs to be addressed.”</p>	<p>“The population was different 80 years later and this play was shown in New York, not Virginia so it wouldn’t interest many.”</p>	<p>“Fact 1 does not provide evidence for why the authors wrote the play because the population they are writing for is the 1930s, not the people of the 1850s.”</p>
Level	Percentage										
Basic	84%										
Emergent	8%										
Proficient	8%										
<p><i>Thanksgiving</i></p> <p>The painting <i>The First Thanksgiving</i> 1621 helps historians understand the relationship between the Wampanoag Indians and the Pilgrim settlers in 1621. Do you agree or disagree? Briefly support your answer.</p>  <table><tr><th>Level</th><th>Percentage</th></tr><tr><td>Basic</td><td>70%</td></tr><tr><td>Emergent</td><td>10%</td></tr><tr><td>Proficient</td><td>20%</td></tr></table>	Level	Percentage	Basic	70%	Emergent	10%	Proficient	20%	<p>“I agree because in 1621 this was most likely what the relationship between the Wampanoag Indians and the Pilgrim settlers were like in 1621 until the settlers became fully independent and knew how to work the land.”</p>	<p>“The painting gives a biased example of the relationship between the Wampanoag Indians and the pilgrim. The artist is definitely not an Indian, he’s probably white. Therefore he’s painting from his perspective and not the Indians and what actually may have looked like.”</p>	<p>“The painting was painted in 1932 for an event that happened in 1621. It is not a primary source. The painting, unless supported by written first-hand accounts, is a account of how a artist in 1932 felt the first Thanksgiving happened.”</p>
Level	Percentage										
Basic	70%										
Emergent	10%										
Proficient	20%										

This table provides examples of the range of answers provided in response to each of the assessment tasks. We used a three-point rubric to score responses: “Basic” (zero points) if the

answer was off base and bore no relation to the competency being measured; “Emergent” (one point) if the answer showed inklings of proficiency; and “Proficient” (two points) if the answer demonstrated understanding. Across the three assessments were six questions (one for the Thanksgiving painting, two on the Philippine-American War documents, and three on the John Brown playbill), resulting in a possible total score of twelve points. SOURCES: For task 1, Testimony of Richard T. O’Brien, U.S. Congress, Senate, Committee on the Philippines, *Affairs in the Philippines: Hearings before the Committee on the Philippines of the United States Senate*, 57 Cong., 1 sess., April 2, 1902, pp. 2549–51. For task 2, George Goldschmidt, “‘Battle Hymn’ a New Play about John Brown of Harpers Ferry by Michael Blankfort and Michael Gold at the Experimental Theatre,” ca. 1936–1941, illustration, *Library of Congress Prints and Photographs Online Catalog*, <http://www.loc.gov/pictures/item/98516478/>. For task 3, Jean Leon Gerome Ferris, *The First Thanksgiving 1621*, ca. 1912–1930, painting, *Library of Congress Prints and Photographs Online Catalog*, <http://www.loc.gov/pictures/item/2001699850/>.

argued, “This [letter] does not show public opinion but one man’s rude, unethical, and racist opinion of people.” Another wrote:

[Funston’s letter] also shows how Americans opposed the war in the Philippines because of the racist views supporters had. Colonel Frederick Funston dismisses opposition by saying that they are “educated, however, about the same way a parrot is” and that they deserve strict discipline to get them in order. Thus, this shows that Americans opposed a racist war.

Only six students out of forty-nine were able to see how the publication of Funston’s letter might provide evidence of opposition to the war. (See figure 5.)

Assessing the Future

These results give us pause. If a required survey course is the only history that students are exposed to during college, what ways of thinking do we want them to master? How can we make sure that students develop such ways of thinking? These questions become sharper still when applied to majors. Unlike their peers in computer science or engineering, the vast majority of history majors will not pursue history as a profession but will go into law or finance or any one of a number of professions. Historians have long claimed that historical study teaches critical thinking. Our results suggest that this may not occur by osmosis. Might a more direct approach be necessary?⁹

To ensure that students develop the reasoning skills central to the discipline, we need new tools to gauge their learning. We do not labor under the assumption that our exercises have solved the problems of history assessment. Our tasks are open to numerous challenges, particularly in their failure to exhaust the wide range and richness of historical thinking. At the same time, we believe that for the field to progress, abstract goals must be given concrete form. We agree with the AHA History Tuning Project’s call for students to “contextualize information.” But what does this look like, and how can we find out if students are learning to do it? Our tasks embody one possible form that brief assessments might take. They provide concrete points of reference that ground department-wide collaboration in ways that abstract goal statements do not.¹⁰

⁹ Paul Sturtevant, “History Is Not a Useless Major: Fighting Myths with Data,” *Perspectives on History*, April 2017, *American Historical Association*, <https://www.historians.org/publications-and-directories/perspectives-on-history/april-2017/history-is-not-a-useless-major-fighting-myths-with-data#>.

¹⁰ We have presented here only a sample of the tasks we have created. Other tasks gauge chronological reasoning by asking students to put two historical documents in temporal order using only the content of the documents

New assessments are a start, but they are insufficient by themselves. A collaborative effort to explore new directions in assessment practice must be organized. Our tasks are best understood as formative assessments rather than end-of-course tests. In the assessment literature, formative assessment is distinguished from end-of-course assessment by its purpose: to inform teaching, not to give students a grade. Formative assessment provides a window into student thinking. Moreover, it gives students feedback on whether they are on track to master course content. Rather than waiting to see what students have learned on a final exam, formative assessment allows us to gauge student learning more frequently and tailor instruction more precisely. Instructors can slow down and revisit concepts that students find challenging or pick up the pace on material that students master quickly.¹¹

Formative assessment is rare in the college history classroom. It does not have to be. On the first day of class, instructors could take five minutes and have students complete the task using *The First Thanksgiving 1621*. Rather than grade responses, instructors could use the task as an entry into a conversation about the evaluation of evidence. Alternatively, instructors could collect student responses and quickly scan them to get a better sense of the beliefs students bring to class. The next session could begin with a discussion of evaluating evidence based on representative student responses.¹²

Along with Harvard University's Eric Mazur, the Nobel Laureate Carl Wieman has pioneered the use of clickers (a type of audience response system) to assess student understanding in college science classes. Wieman has shown how instructors can obtain immediate feedback about student thinking by having students respond to prompts he projects from the podium. Nothing is stopping us from doing something similar. Instructors could display one of our tasks and show typical responses, asking students to select which one is best and explain why in small groups. These responses would provide instructors with instant feedback about student understanding instead of assuming that what is second nature to historians is second nature to students.¹³

Student responses also provide opportunities for departmental collaboration. We observed collaboration of this sort at the high school level when we worked with a department that met monthly to discuss student work. At each meeting, teachers reviewed student responses to our exercises and discussed how well students grasped aspects of historical thinking. Over the course of a year, teachers shared strategies for integrating assessments into their courses and developed a shared set of expectations for student learning.¹⁴

or ask students to reason about the strengths and limitations of historical documents as evidence of the past. Still others require students to make connections between seemingly unconnected events across time. To view all of our tasks, visit "Beyond the Bubble," *Stanford History Education Group*, <https://sheg.stanford.edu/history-assessments>. "AHA History Tuning Project."

¹¹ Paul Black and Dylan Wiliam, "Assessment and Classroom Learning," *Assessment in Education: Principles, Policy & Practice*, 5 (no. 1, 1998), 7–74.

¹² Ferris, *First Thanksgiving 1621*.

¹³ Catherine Crouch and Eric Mazur, "Peer Instruction: Ten Years of Experience and Results," *American Journal of Physics*, 69 (Sept. 2001), 970–77; Louis Deslauriers, Ellen Schelew, and Carl Wieman, "Improved Learning in a Large-Enrollment Physics Class," *Science*, May 13, 2011, pp. 862–64. Lendol Calder, "Uncoverage: Toward a Signature Pedagogy for the History Survey," *Journal of American History*, 92 (March 2006), 1358–70; Daniel Immerwahr, "The Fact/Narrative Distinction and Student Examinations in History," *History Teacher*, 41 (Feb. 2008), 199–205; David Pace, "The Amateur in the Operating Room: History and the Scholarship of Teaching and Learning," *American Historical Review*, 109 (Oct. 2004), 1171–92; Sam Wineburg, "Teaching the Mind Good Habits," *Chronicle of Higher Education*, April 11, 2003, <http://chronicle.com/weekly/v49/i31/31b02001.htm>.

¹⁴ Breakstone, "History Assessments of Thinking."

The study of history should be a mind-altering encounter that leaves one forever unable to consider the social world without asking questions about where a claim comes from, who is making it, and how time and place shape human behavior. If the major is to succeed in fulfilling this mind-altering mission, historians cannot be resigned “to suck at assessment.” There may be disagreements about how to define the major, but we doubt that any readers of this article would celebrate the fact that most students ignored the date of a document or failed to consider the context in which it was created. As Anne Hyde noted, the assessment train is barreling ahead. If historians do not create assessments that capture the unique aspects of the discipline, others will come in with their one-size-fits-all tool kit and do the job for them.¹⁵

That would really suck.

¹⁵ Hyde, “Five Reasons History Professors Suck at Assessment,” 1106.