Policy Report

Improving the Measurement of Civic Learning

Madeline J. Goodman, Andrew R. Weiss,
Jesse R. Sparks and Kelsey D. Dreier

THE ETS CENTER FOR RESEARCH ON HUMAN CAPITAL AND EDUCATION

ETS

Table of Contents

Preface

In a recent interview about the war in Ukraine and the future of democracy, the historian Timothy Snyder proclaimed: "The fact that we [in the United States] have a democracy at all is kind of remarkable."[1] This statement serves as a reminder of the perils that democratic governments and societies have faced over time. In past moments – the period of the Civil War and Reconstruction, as well as the two world wars of the past century come to mind – the future of American democracy, and democracy itself, was far from a foregone conclusion. Democracy is more fragile, and more dependent on the continued choices we need to make individually and collectively, than many presume. Among other things, democracy is contingent upon the active participation of citizens and a citizenry educated in knowledge of government and history as well as the skills of deliberation, cooperation, and inquiry. *Improving the Measurement of Civic Learning* explores how well we are preparing our students to become well-informed, productively engaged citizens, and the role that assessments, particularly large-scale assessments, can play in improving the conceptualization and measurement of civic learning for our nation's students.

The first part of the report focuses on data from the Nation's Report Card and explores eighth grade student performance on the National Assessment of Educational Progress (NAEP) civics and U.S. history assessments. While journalists and researchers often cite that U.S. students (and adults) lack knowledge about how our government functions or key aspects of the nation's past, authors of this report assert that this is perhaps not the sole—or even the fundamental—concern. The percentages of students that perform below the *NAEP Proficient* level—75 percent of students fall below this level in civics and 85 percent in U.S. history—is undoubtably alarming. These data highlight that many of the nation's students are struggling to understand and apply knowledge and skills in civics and U.S. history. They also expose the extent to which our educational institutions and policy makers have consistently failed to prioritize teaching and learning in these subjects. Despite a frequent lament that students and young adults do not have the critical thinking and problem skills they need, we continue to devalue learning in some of the very domains where such skills are acquired and honed. The skills embedded in civics and U.S. history—the ability to evaluate critically texts of all kinds, to engage in civil deliberation with others, to understand other perspectives and grasp the process of change over time in our society and democracy—are skills we must explicitly teach students and are generally taught through a disciplinary lens.[II] As the report asserts, we need to align standards, curriculum, and assessment frameworks to target the teaching and learning of these critical thinking and deliberative skills so that our students are learning core competencies they need to fully participate in our democracy.

In contextualizing the performance results, the report authors also reflect on the fundamental role that literacy plays in acquiring and demonstrating proficiency in disciplinary domains such as civics and U.S. history. Large-scale assessment results from NAEP, the Programme for International Student *Assessment* (*PISA*), and the Programme for the International Assessment

of Adult Competencies (PIAAC) reveal that too many of our nation's middle and high school students and young adults do not demonstrate the levels of literacy necessary to perform tasks that involve critical reading and evaluation of printed or digital texts.[III] It may be of little surprise then that a large proportion of eighth graders struggle to demonstrate they are at the *NAEP Proficient* level in civics and U.S history. As others have argued,[IV] higher order skills such as critical thinking, problem solving, and evaluation of text and nontext sources are built upon foundational literacy skills and tied to content knowledge: if students struggle to read, they will likely find it challenging to acquire higher order skills in domains such as civics and U.S. history.

Policy makers, educators, and the public in general will appreciate that the report offers a nuanced view of the 2018 NAEP data—the only source of data on students' knowledge and skills in civics and U.S. history based on a nationally representative sample. The authors use a regression analysis to explore the role that demographic factors, socioeconomic status (SES), what is often referred to as opportunity to learn (OTL; in this study, is defined as teacher preparation and classroom practices), and students' dispositions play in achievement outcomes. The analyses here reveal that SES and demographic factors such as race and English language learning status—which are complexly entangled—account for approximately 30 percent of the variance in students' scores on the 2018 civics and U.S. history assessments. Adding OTL factors as currently measured by NAEP into the regression model resulted in only incremental increases in explained variance in scores. This suggests that issues of equity that impact students outside of school, the SES of their families and neighborhoods, for example, continue to exert a critical influence on the quality of education students receive. In addition, the results indicate that part of the reason students struggle to acquire the knowledge and skills they need to be more proficient in these areas may be due to lack of adequate training and preparation of teachers. If we see civic learning as only knowledge acquisition and neglect the core critical thinking skills that are embedded in this domain, then we will not adequately train our teachers to impart these higher order skills, and students will struggle to learn them.

The second part of the report digs even deeper into how we can better conceptualize and measure civic participatory skills and historical thinking skills alongside core knowledge about our government and history. In this section of the report, the authors discuss recent work to define civic participatory and historical thinking skills more clearly, and the implications this has for improving assessments. As just one example, the authors not only describe a skill such as "historical causation" as the "ability to identify, analyze, and evaluate multiple cause-and-effect relationships in a historical context, distinguishing between the long-term and proximate," they go further to provide concrete examples of tasks and questions that target the measurement of this skill. What the authors demonstrate here is something that those in the assessment community—and educators on the ground—know very well: the more complex the skill or knowledge we wish to impart, the harder it is to measure that skill. The description of new task-types and sample tasks provide tangible examples for educators and

educational policy makers concerned with a renewed, rigorous focus on improving the teaching of core skills in civic learning and integrating these with the acquisition of knowledge in these domains.

Given the importance of NAEP as the only comparable monitor of student progress across the nation in civics and U.S. history, as well as the critical role that a well-designed assessment can play in the learning process, the intention expressed here to improve the measurement of NAEP in these areas is both timely and essential. The underlying message of the report is clear: To strengthen our democracy and civil society, we need to rethink how we educate our children and how we prioritize critical thinking both in and outside of the classroom.

Madeline Goodman and Irwin Kirsch
Center for Research on Human Capital and Education

Notes

[I] Sean Illing, "Ukraine and the problem of "futurelessness": Historian Timothy Snyder on the war in Ukraine and the future of democracy." *Vox*. June 21, 2022. https://www.vox.com/2022/6/21/23165718/vox-conversations-ukraine-russia-timothy-snyder-democracy.

[II] Susan R. Goldman, M. Anne Britt, Willard Brown, Gayle Cribb, MariAnne George, Cynthia Greenleaf, Carol D. Lee, Cynthia Shanahan, and Project READI, "Disciplinary Literacies and Learning to Read for Understanding: A Conceptual Framework for Disciplinary Literacy," *Educational Psychologist* 51, no. 2 (2016): 219–46, https://www.researchgate.net/profile/Susan-Goldman/publication/301795083_Disciplinary_Literacies_and_Learning_to_Read_for_Understanding_A_Conceptual_Framework_for_Disciplinary_Literacy/links/58d8450092851c44d4ad2dc7/Disciplinary-Literacies-and-Learning-to-Read-for-Understanding-A-Conceptual-Framework-for-Disciplinary-Literacy.pdf

[III] OECD, *PISA 2018 Results (Volume I): What Students Know and Can Do* (Paris: OECD Publishing, 2019), https://doi.org/10.1787/5f07c754-en; Bobby D. Rampey, Robert Finnegan, Madeline Goodman, Leyla Mohadjer, Tom Krenzke, Jacquie Hogan, and Stephen Provasnik, *Skills of U.S. Unemployed, Young, and Older Adults in Sharper Focus: Results from the Program for the International Assessment of Adult Competencies (PIAAC) 2012/2014. First Look*, NCES 2016-039 (Washington, DC: National Center for Education Statistics, 2016), https://nces.ed.gov/pubs2016/2016039.pdf; for NAEP reading results at Grade 8, see "NAEP Report Card: Reading. National Achievement-Level Results," The Nation's Report Card, accessed February 8, 2023,

https://www.nationsreportcard.gov/reading/nation/achievement/?grade=8; for NAEP results at grade 12, see "NAEP Report Card: Reading. Explore Results for the 2022 NAEP Reading Assessment," The Nation's Report Card, National Center for Education Statistics, accessed February 8, 2023, https://www.nationsreportcard.gov/reading/?grade=12.

[IV] Goldman et al., "Disciplinary Literacies."

Acknowledgements

Introduction

> "
>
> *Liberty, opportunity, and mutual respect are not self-evident or self-perpetuating. They must be carefully taught or else opposing values—authoritarianism, plutocracy, intolerance, bigotry, and hatred—will dominate our societies.[1]*
>
> —*Amy Guttman*
> *President Emeritus, University of Pennsylvania*
> "

The association between education and democracy, alluded to in the statement by political theorist Amy Guttman, dates as far back as Greece circa the 5th century BCE. While Greek society and political institutions were by no means perfect models of democracy, the Greeks forged a critical, enduring link between democratic, civil society and the need to inculcate democratic values through education. In the United States, the idea that the schools should serve as the training ground of democracy extends back to the nation's founding and the beginnings of public schooling in the antebellum period. Among the many, often conflicting, goals of public education have been the strengthening of democracy and the training of youth to take on the responsibilities of citizenship.[2] Focus on this goal has waxed and waned over the years, frequently rising to prominence during periods of domestic turmoil, such as in the post-civil war period, and often spilling over from the confines of the education community into a broader public discourse. Calls for a renewed commitment to civic education have come at a steady pace since the 1983 release of *A Nation at Risk*, a report aimed at the highest levels of government claiming that too many Americans lacked the skills and motivation to sustain democracy. Authors of the report posited that "for our country to function, citizens must be able to reach some common understandings on complex issues, often on short notice and on the basis of conflicting or incomplete evidence. Education helps form these common understandings."[3] This emphasis was followed over the years by the periodic release of civics standards, curriculum frameworks, re-conceptualizations of the domain, and renewed calls for action, the latter of which suggests that despite a great deal of effort by many educators, scholars, and philanthropic organizations concerned with preserving democracy, the warnings went largely unheeded by the nation's education policymakers.[4]

The purpose of this report is to advance the discussion of how we understand civic learning and the role large-scale assessments can and should play in this process. The intention is to provide policy makers, educators, and those interested in civic learning with insights into what current large-scale, national data on student outcomes in civics and U.S. history tell us about students' knowledge and skills. In addition, the report explores a body of work that seeks to improve the measurement of core skills and competencies embodied in civic learning by

applying advancements in measurement and learning sciences to the domains of civics and U.S. history. We relied on current efforts to improve civic learning by defining this domain as one with several, interrelated competencies that range from a knowledge and understanding of history and government to the ability to apply critical thinking and inquiry skills to that knowledge.[5] It is clear from recent initiatives in civic learning that both instruction and assessment must go well beyond the rote learning of isolated facts about U.S. history and how our government functions to include the development of critical thinking, reasoning, and participatory skills in order to raise the level of civic discourse and meaningful engagement in our democracy.[6] The paper explores these themes by (1) setting a context for why a more comprehensive understanding of critical thinking skills forms a central component of civics and U.S. history and deepening our understanding of what it means to be "literate" in these domains; (2) investigating the 2018 grade 8 data from the National Assessment of Educational Progress (NAEP) in civics and U.S. history regarding what eighth graders know and can do and how student performance is associated with demographic, socioeconomic status (SES), opportunity to learn (OTL) factors, and students' attitudes toward these subject areas; and (3) exploring how large-scale assessment can be designed to measure more effectively the skills critical to civic understanding and engagement. Finally, we will discuss the implications of designing assessments that provide better data that can improve civic learning.

Following the example of many educators and scholars in the field, we use the term "civic learning" rather than the more standard "civic education" to acknowledge the broad range of practices and activities that comprise effective education for democracy and the fact that some of these activities take place outside the classroom through forms of service learning, extracurricular activities, civic action, and other types of community engagement. We use the term in full recognition that, in general, the farther students move from classroom learning to civic action, the more challenging it becomes to assess their skills and abilities using traditional approaches to educational measurement, particularly in the context of a large-scale assessment.[7]

This report also deliberately and explicitly includes history as a cornerstone of civic learning. There is a diversity of opinion on the role of history in civic learning, ranging from the belief that learning civics requires a thorough study of history to one in which history plays a relatively minor supporting role to political science, current affairs, and civic action. The reasoning in support of history's importance is persuasive, however. As the authors of a recent National Academy of Education report noted, "Civic reasoning and discourse inevitably involve how members of a society see themselves, and are also inevitably related to understanding our history as a nation, and what that history reveals about the who and what of the United States."[8] Although principles related to civic participatory skills such as improved communication, conflict management, problem-solving, and reasoning can be learned and supported by many subject areas , the knowledge and skills gained through the study of history are central. There is a shared concern with a base of knowledge about how our constitutional democracy was established and changed over time in response to political,

social, economic, and technological forces. In addition, history is a discipline focused on inquiry, on *how* we know what we know about the past. Key skills that are part of the domain of history—what we term here *historical thinking skills*—include understanding change over time, comprehending the role of bias and perspective, reading and critically evaluating historical sources, and synthesizing information and formulating arguments. Many of these inquiry skills are also central to the exercise of civic learning, as underscored in recent frameworks.[9] One might go so far as to suggest that the skill of perspective taking or even the empathy we develop through the study of history can have a significant positive impact on the quality of our civic discourse, as aspects of these skills are often incorporated into broad definitions of social and emotional learning competencies seen as central to civic learning.[10] Noting "civic and political dysfunction" in the United States, it makes sense that the recently released *Roadmap to Educating for American Democracy* calls for "achieving excellence in history and civic education in support of civic strength."[11]

Civic Learning and Literacy

As James Pellegrino, Louis DiBello, and Susan Goldman have pointed out in their research, recent calls for new standards and assessments in various disciplines proceed with an understanding that "knowledge in the form of memorized facts and rote procedures is inadequate to support flexible, creative, and innovative problem solving..."[12] These scholars are in fact pointing to a broader understanding of literacy that is particularly relevant to the discussion of subject area knowledge in such domains as history, civics, and the sciences. Part of the challenge of measuring students' knowledge and skills in these domains is that the domains themselves are strongly influenced by the ability to read and decipher textual information. Courts around the country have recently heard arguments and weighed in regarding the strong tie between literacy, public education, and democratic citizenship.[13] As a Michigan court succinctly claimed in a recent ruling: "without the literacy provided by a basic minimum education, it is impossible to participate in our democracy."[14]

The ability to decode and comprehend texts, that is, to exhibit essential literacy skills, is inextricably tied to other higher order skills associated with domain-specific literacy in civics and U.S. history and with the kinds of skills and knowledge we want our students to possess as they move through the K-12 educational system and beyond.[15] Yet it is important to underscore that the need for literacy extends beyond deciphering text-based material. For example, students learning civics and U.S. history should be able to decode meaning from myriad non text-based sources, such as video, newsreels, audio interviews, music, photography, maps, and compilations of data, to name just a few. Admittedly, the capacity to do so may be tied to literacy or "reading" skills, but the extent to which the one informs the other is difficult to distill and measure. The cross-subject *NAEP Proficient* achievement level is described as performance that demonstrates "solid academic performance... [and] competency over challenging subject matter."[16] To the extent that this performance involves the use of higher order skills in civics and U.S. history, such as problem-solving, critical

thinking and the ability to evaluate sources and arguments, we need to acknowledge how these are built upon—and extensions of—essential literacy skills.[17] We must also appreciate their domain-specific qualities; one does not think critically or solve problems about nothing. Thinking critically requires one to apply skills (some of which may be domain-specific) to a body of knowledge.[18] As scholars have noted, domains of knowledge embody epistemic beliefs about how particular knowledge is constructed. In their research, Goldman and colleagues see epistemology as "providing purpose and motivation to the ways in which inquiry was conducted, the reasoning principles that were invoked; and the forms in which information was represented, expressed, examined, and critiqued, and negotiated in and through oral and written discourse and language structures."[19]

The implications of acknowledging the domain-specific nature of knowledge and literacy in terms of the present report are twofold. First, the performance of students on large scale assessments of civic and U.S. history may in part reflect low levels of basic literacy skills in our students and the interconnectedness of domain knowledge and use of higher order skills. To better inform educational practice, large-scale assessments in these domains may need to do more to measure students' capabilities at the lower end of the performance distribution to help disentangle whether and how low levels of foundational literacy skills impede the ability of students to demonstrate higher order skills. Second, standardized summative assessments can likely improve their ability to measure discipline-based higher order inquiry skills in civics and U.S. history. Any such improvement will align assessments with how experts in the field currently conceptualize these domains. Later sections of this paper will explore how we can enhance the measurement of participatory and critical thinking skills in civics and U.S. history and will discuss the interconnectedness of knowledge and skills in civics and U.S. history.

NAEP 2018 Grade 8 Civics and U.S. History Results

Rare is the popular or scholarly work about the current state of our constitutional democracy that does not use NAEP results to argue that our students are ill prepared in terms of their knowledge and skills in civics and U.S. history. For example, NAEP data frequently provides the evidence for claims, such as those made in a recent National Academy of Education report on civic reasoning and discourse, that we need to better "prepare a new generation of young people to take up the mantle of democratic participation and decision making."[20] These large scale national assessments have come to play an important role in the debate about whether our students are acquiring the knowledge and skills they need to actively participate in our democracy.

NAEP civics and U.S. history data is collected every four years at the national level and provides the only assessment of these domains based on a nationally representative sample of students. In 2018, approximately 13,400 eighth graders from 780 schools took the civics assessment and 16,400 eighth graders from 780 schools took the U.S. history assessment. The

assessments are written to frameworks developed in the early-to-mid 1990s. From 1994 to 2010 in U.S. history and 1998 to 2010 in civics, assessments were given at grades 4, 8, and 12. From 2010 forward, the assessments have been given only at grade 8.

Some additional information about the NAEP assessments may be helpful for readers. NAEP is designed to provide results on the performance of student *groups* (e.g., gender, race/ethnicity, eligibility status for the National School Lunch Program [NSLP]), as well as to allow for group comparisons over time; it is not designed to provide individual student scores.[21] The assessments use selected-response and constructed-response questions to measure student performance. In 2018, the civics and U.S. history assessments transitioned from paper-and-pencil format to a digital format; however statistical linking analyses were performed to allow for trend comparison across assessment modes.[22] In addition to gathering information on students' cognitive skills in these subjects, data are collected from student, teacher, and school administrator responses to survey questions about learning attitudes, classroom practices, and school environments. NAEP teacher and school data, used in the regression analyses reported here, are linked to students who took the NAEP assessment in the respective subject area: therefore, it is possible to discuss self-reported teacher training and classroom activities in relation to student group level performance on the NAEP civics and U.S. history assessments.

This section will explore student performance in civics and U.S. history by looking at descriptive data regarding overall score trends, trends for the NAEP achievement-level results, as well as score gaps among some student groups (e.g., racial/ethnic groups, gender, SES indicators). Next, we will provide results and discussion around a series of regression analyses to explore how factors related to demography, SES, OTL, and student attitudes are associated with performance in NAEP civics and U.S. history. As educational researchers have repeatedly claimed, the decentralized nature of education in the United States results in an uneven distribution of quality education.[23] Moreover, factors related to SES (often confounded by race, ethnicity, language, and geography) powerfully shape students' experiences in ways that differentially affect student achievement outcomes.[24] The analyses here attempts to tease out from the only nationally representative sample of data on students' knowledge, skills, and practices how issues related to equity and opportunity to learn are associated—albeit in complex ways—to student outcomes in civics and U.S. history.[25]

Performance on NAEP is measured using two metrics: (1) an average scale score on either a 0–500 (U.S. history) or 0–300 (civics) scale;[26] and (2) a percentage of students performing at three achievement levels: *NAEP Basic, NAEP Proficient,* and *NAEP Advanced*. While below *NAEP Basic* is not an official achievement level, NAEP does report the percentage of students falling below that level. NAEP achievement levels, set by the National Assessment Governing Board, are essentially bands of performance on the subject scale determined by the collective judgments of a broadly representative panel of teachers, educators, and members of the general public. NAEP achievement levels are not intended to correspond to grade levels.[27]

That is, performing at the *NAEP Proficient* level does not denote performance at grade level and performance at the *NAEP Basic* level does not denote performance below grade level. In the discussion below, any mention of differences in scores or percentages are based on statistical significance set at the .05 level, with appropriate adjustments for multiple comparisons.

In a broad sense, across all subject areas assessed by NAEP, students who perform at the *NAEP Basic* level demonstrate *partial* mastery of skills, while those at the *NAEP Proficient* level demonstrate "solid academic performance...[and] competency over challenging matter, including subject matter knowledge, application of such knowledge to real-world situations, and analytical skills appropriate to the subject matter."[28] In civics, students performing at the *NAEP Proficient* level should, among other things, understand and be able to explain the purposes that government serves, have a good understanding of the difference between government and civil society, recognize the distinction between American ideals and reality, explain ways that citizens can influence the government, and give simple interpretations of information from non text sources. Students performing at this level in U.S. history should be able to explain the significance of people, places, events, ideas, and documents; recognize the connection between people and events within historical contexts; and understand and be able to explain the opportunities, perspectives, and challenges associated with a diverse cultural population. In addition, they should incorporate geographic, economic, social, and political considerations in their understanding of events and be able to communicate ideas about historical themes while citing evidence from primary and secondary sources to support their conclusions. In other words, students who perform at the *NAEP Proficient* level demonstrate an ability to use higher order thinking skills (analysis, evaluation, interpretation) in relation to content in civics and U.S. history, while those performing at the *NAEP Basic* level or below will likely struggle to demonstrate the kinds of skills associated with higher order thinking in these domains.[29] For a detailed description of NAEP achievement levels in civics and U.S. history, see Appendices A and B.

Figure 1: Overall NAEP Scale Score Trend in Eighth Grade Civics and U.S. History
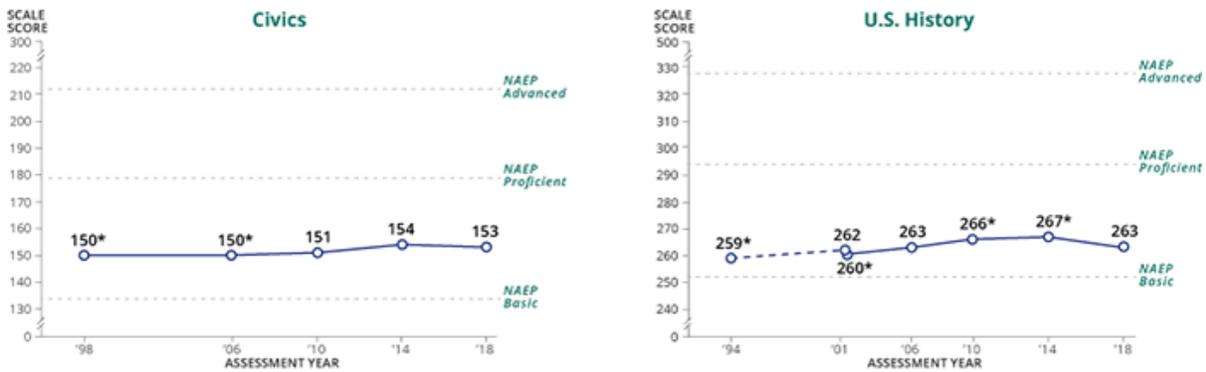
* Significantly different (*p* < .05) from 2018.

Note: The NAEP civics scale ranges from 0 to 300 and the U.S. history scale ranges from 0 to 500. Not all apparent differences between estimates are statistically significant.

Source: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), various years, 1998–2018 Civics Assessments and 1994–2018 U.S. History Assessments.

Figure 2: Trend in Eighth Grade NAEP Civics and U.S. History Achievement Level Results

[1] Accommodations not permitted.

* Significantly different (*p* < .05) from 2018.

Note: NAEP achievement levels are to be used on a trial basis and should be interpreted and used with caution.

Source: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), various years, 1998–2018 Civics Assessments and 1994–2018 U.S. History Assessments.

Data from the 2018 civics assessment in Figure 1 show that the average score for eighth graders nationally remained virtually unchanged from the previous assessment year (2014) but was higher than in the first year the assessment was administered (1998). In U.S. history, the average score decreased by four points compared to 2014 but was higher than in the first assessment year (1994). In 2018, 76 percent of students performed below the *NAEP Proficient* level in civics (49 percent performed at the *NAEP Basic* level and 27 percent performed below the *NAEP Basic* level), and there was no statistically significant change in this percentage compared to any of the other years the assessment was administered (see Figure 2). In U.S. history, 85 percent of eighth graders performed below the *NAEP Proficient* level (50 percent performed at the *NAEP Basic* level and 34 percent below the *NAEP Basic* level). While the percentage of students performing at or above the *NAEP Proficient* level in 2018 decreased compared to the previous assessment year (2014), there was no significant change in the percentage of students at this level compared to 1994 for U.S. history.

While not shown here, gaps in performance among select student groups have been persistent over time in both subjects across key demographic variables (e.g., race/ethnicity) as well as SES variables (parental education and NSLP eligibility are the main SES-related variables highlighted in initial release reports from the National Center for Education Statistics [NCES]).[30] While score gap sizes are not directly comparable across NAEP subjects, it should be noted that gaps in performance between racial/ethnic and SES groups are observed in every subject area and grade level and likely reflect broad equity concerns that are not subject-specific. As many scholars have recognized, because race/ethnicity is intrinsically tied to SES in the United States, it is likely that performance results by race reflect differences in access to opportunity in terms of school quality, neighborhood stability and safety, and availability of home and school resources that support learning for some racial/ethnic groups vis-à-vis others.[31] One notable change in such gaps over time occurred in civics, where, compared to 1998, the White–Hispanic student score gap narrowed, in large part due to greater increases in the performance of Hispanic vs. White eighth grade students.[32] Nonetheless, sizable gaps among racial/ethnic and SES groups remain in both subjects.

There is a need to gain a more nuanced understanding of the results than can be attained by the aggregated data alone. A fuller understanding of the issue would use available contextual information based on student, teacher, and school reports to interpret the aggregate results. The next section addresses this issue.

Analyses of Factors Related to Performance

As many scholars and policy makers have pointed out, advancing teaching and learning in civics and U.S. history involves not only improving content instruction and assessment but attending to key equity concerns.[33] Variations in access to quality education in civic learning have—not surprisingly—been identified as being associated with factors related to access to quality schools, such as living in a high-income neighborhood and attending a school with more resources.[34] This dictates that we examine available data to help us understand what groups of students know and can do, but also how student achievement is associated with factors related to demography, SES, and aspects of OTL both in and outside the classroom. To gain a more nuanced understanding of the NAEP civics and U.S. history results than can be derived from aggregated data such as average scale scores and NAEP achievement level performance, and to help us determine how that data can be either used to inform policy or be improved to better do so, we performed a series of regression analyses using the 2018 data.

The analyses below use the 2018 civics and U.S. history scale scores as the outcome variable in a series of regression analyses. Hierarchical multiple linear regression was utilized to better understand and isolate the impact of both individual variables and indices related to student demographics, OTL, and select student intra/interpersonal factors. Indices are clusters of correlated variables that measure underlying constructs or concepts such as enjoyment of the

subject, motivation, engagement, etc. We ran preliminary analyses to estimate the relationships between scale scores and variables/indices associated with SES, OTL, and student intra/interpersonal factors related to core skills in civics and U.S. history. We then selected variables/indices for our regression models (see Table 1). We explored the influence of key demographic characteristics (Model 1); SES factors (Model 2); a cluster of OTL variables related to teacher background (Model 3) and teacher classroom emphasis on particular skills associated with civics learning (Model 4); and, finally, selected student intra/interpersonal indices that reflect students' interest, motivation, and confidence in these domains (Model 5). The research here is exploratory and examines the associations between student scores at the group level and a cluster of data gathered from the student, teacher, and school questionnaires that researchers have previously identified as important to academic success. These models do not assess whether the explored relationships are causal. In addition, performing these regression analyses allows us to tease apart the extent to which more persistent factors, such as students' demographic and SES status, are associated with some students' ability to demonstrate their knowledge and skills on large-scale assessments such as NAEP and should therefore be considered alongside descriptive data of students' performance results. The intention here is to broaden the discussion of the factors related to achievement in civics and U.S. history and suggest avenues for future research in terms of classroom practice, teacher background and preparation, and the measurement of variables included in our models. For more detailed information regarding methodology used in the analyses, see Appendix C.

Table 1: Variables Used in Regression Analysis

CIVICS	U.S. HISTORY
DEMOGRAPHICS (MODEL 1) • Gender • Race/ethnicity • English language proficiency	**DEMOGRAPHICS (MODEL 1)** • Gender • Race/ethnicity • English language proficiency
SOCIOECONOMIC STATUS VARIABLES (MODEL 2) • Books in the Home • Parent's Education • School NSLP Status • Student NSLP Status	**SOCIOECONOMIC STATUS VARIABLES (MODEL 2)** • Books in the Home • Parent's Education • School NSLP Status • Student NSLP Status
OPPORTUNITY TO LEARN VARIABLES	
TEACHER: BACKGROUND/PREPARATION (MODEL 3) • Teaching Experience • Professional Development • Political Science Degree • Civics Teaching Role • Teacher Satisfaction	**TEACHER: BACKGROUND/PREPARATION (MODEL 3)** • Teaching Experience • Professional Development • History Degree • History Teaching Role • Teacher Satisfaction
TEACHER: CLASSROOM EMPHASIS (MODEL 4) • Students writing opinions • Students taking part in role-play/mock-trial • Students using computer to create reports • Emphasis on discussing political process • Emphasis on taking position on issues	**TEACHER: CLASSROOM EMPHASIS (MODEL 4)** • Students use Internet for evidence • Assignment with multiple sources • Assess long written responses
STUDENT INTERPERSONAL/INTRAPERSONAL INDICATORS (MODEL 5) • Civics interest/enjoyment index • Civics community engagement index • Importance of attention to government • Importance of participation in government	**STUDENT INTERPERSONAL/INTRAPERSONAL INDICATORS (MODEL 5)** • U.S. history interest/enjoyment index • Students' confidence in U.S. history knowledge/skills index • Students' perspective taking in historical thinking index

NOTE: For further details on regression variables, see Appendix C, Tables C2 and C3.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2018 Civics and U.S. History Assessments.

Table 2: Explained Variance (R^2) by Model in NAEP Eighth Grade Civics and U.S. History: 2018

VARIANCE	MODEL 1	MODEL 2	MODEL 3	MODEL 4	MODEL 5
CIVICS					
R^2	0.177	0.303	0.308	0.350	0.387
ΔR^2	0.177**	0.126**	0.005	0.042**	0.037**
U.S. HISTORY					
R^2	0.193	0.314	0.322	0.330	0.362
ΔR^2	0.193**	0.121**	0.008	0.008	0.032**

Statistical significance: * $p < .05$, ** $p < .01$.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2018 Civics and U.S. History Assessments.

Table 2 shows the explained variance in scale scores for each model that accounts for factors related to demographics, SES, OTL, and student attitudes across the two subjects. (Note: To explore scale score difference for the individual variables within the various models, see Appendix C, Tables C4 and C5 for civics and U.S. history results). Model 1, which includes gender, race/ethnicity, and English learner (EL) status, explains approximately 18 percent of the variance in student scores in civics and 19 percent of the variance in U.S. history scores. While SES is notably difficult to measure in large-scale assessments of student achievement, it is nonetheless generally accepted practice to use available proxy variables, such as parental education level, number of books in the home, NSLP eligibility, and percentage of students in the school who are NSLP eligible.[35] When we add these SES-related variables (Model 2), the explained student score variance increases from 18 percent to 30 percent in civics and from 19 to 31 percent in U.S. history. This means that approximately 70 percent of the variance in students' scores is attributable to factors unrelated to those captured by the demographic and SES measures available in the NAEP data. Of the variables included in Model 2, the number of books in the home, a long-standing SES proxy used in international and national assessments, has a relatively large impact on scores in both subject areas. In 2018, students who reported having more than 100 books in their home scored 12 points higher in civics and U.S. history as compared to students who reported having 0–10 books in the home, when controlling for demographic factors and our other SES measures (see Appendix C, Tables C4 and C5).

The addition of OTL variables to the models was done in two phases in order to explore how OTL, which research has shown is often related in complex ways to race and SES, is associated with students' scores when we control for demographic and SES factors.[36] We operationally define OTL, a concept that is challenging to delineate and consistently measure,[37] as including (1) a set of factors/variables related to teacher background and (2) a set of factors/variables related primarily to the emphasis teachers place on various learning activities in the classroom (see Table 1). We did this to explore the differential impact of these aspects of OTL. Learning activities as examined here generally refers to those that involve "doing" civics or

using history thinking practices, for example, participating in debates, writing an opinion about a specific issue/problem, interpreting or using primary sources, or having students answer questions using multiple types of sources. Research supports the use of these kinds of classroom activities as having an association with student outcomes.[38]

First, we looked at the change in explained variance when adding teacher background factors (i.e., undergraduate major in political science or U.S. history, years teaching social studies, whether teachers had a primary role teaching U.S. history or civics, teacher satisfaction in their school, and whether teachers had received professional development in the past two years), controlling for factors in Models 1 and 2. Literature on OTL cites these as important elements of teacher preparedness that are related to student performance.[39] Including all these variables together as "teacher background" in Model 3 had very little impact on the explained variance in either subject compared to Model 2.

The second group of OTL factors relates to teacher instructional practices (Model 4).[40] We focused attention on the teacher survey questions we believed were most closely related to core skills associated with civics and U.S. history learning, such as those related to inquiry, historical (critical) thinking, and participatory aspects of civic engagement. While research on OTL in terms of classroom exposure to content and skills in the social sciences is limited, there is general agreement that it can be measured in part by asking teachers what type of emphasis they give to specific subject-related practices in the classroom.[41] At the same time, there are potential problems in relying on teacher-reported data on classroom practices. For example, there is likely variation in the way teachers interpret the use of specific practices, which can complicate the meaning imposed on the aggregated data.[42] After including these factors in Model 4, the explained variance in civics increased from 31 percent in Model 3 to 35 percent in Model 4 and in U.S. history from 32 percent in Model 3 to 33 percent in Model 4. Overall, this suggests that after controlling for student demographics, SES, and teacher background variables, classroom activities largely related to active learning and domain skills in civics and U.S. history had a relatively small impact overall on the variance of student scores on the NAEP civics and U.S. history assessments. To explore this further, we ran additional analyses to isolate the independent effect of the sets of variables in each of our models on student achievement outcomes. The individual effect by subject for each group of variables was not much different from what we observed in the hierarchical linear regression model. In civics and U.S. history, race/ethnicity and gender accounted for 19 percent of variance in scores and SES status for 27 percent of variance in scores. Teacher background accounted for 3 percent of the variance in students' civic scores, and 4 percent of variance in students' U.S. history scores, while teacher emphasis accounted for 5 percent of the variance in students' civics score and 4 percent of variance in U.S. history scores.

While the explanatory power of teacher classroom practices was relatively low, there were some practices that had a stronger association with students' scores after controlling for other factors. In U.S. history, students whose teachers reported they assigned students tasks

involving the use of multiple sources for "more than half of the lessons" scored seven points higher than those whose teachers reported devoting no time to the use of multiple sources during their lessons, when controlling for all other factors in Model 4. In civics, students whose teachers reported having students participate in mock trials or debates four to five times in a year scored 3 points higher than those whose teachers never had them do so, when controlling for all other variables in Model 4.

Our final model (Model 5) included individual variables and indices related to students' engagement with core practices and skills in civics and U.S. history. These types of intra- and interpersonal factors—that is, strategies, motivations, and behaviors that are somewhat distinct from content knowledge and academic skills—have been found to be positively related to students' academic performance, high school completion, persistence in postsecondary education, and long-term outcomes, including their success in the job market.[43] For civics, we included indices that captured students' enjoyment and engagement in the subject, as these reflect the civic behaviors and dispositions that are core to civic engagement and learning. (Note: for a list of the individual questions included in the indices for civics in Regression Model 5, see Appendix C, Table C2). In addition, for the civics model, we included data on an individual student survey question related to the importance of civic participation dispositions (i.e., Can you explain the importance of paying attention to government actions?). The inclusion of these indices/variables in Regression Model 5 increased the explained variance in scores in civics by 4 percentage points compared to Model 4, (from 35 percent in Model 4 to 39 percent in Model 5). That is, after including a wide array of variables to control for demographic factors, SES, OTL, and key student attributes related to civic participation, our final model accounted for 39 percent of the variance in student scores on the 2018 NAEP civics assessment.

In U.S. history, our final regression model included three indices related to students' perceptions of history and their ways of thinking about the past. The questions embedded in these indices were meant to probe students' interest/enjoyment of U.S. history, confidence in their ability to use key historical thinking skills, and conception of themselves as those who approach issues with "a historical mindset."[44] (Note: for a list of the individual questions included in the indices for U.S. history in Regression Model 5, see Appendix C, Table C3.) With these factors included in the final model, the explained variance in students' scores increased 3 percentage points (from 33 percent in Model 4 to 36 percent in Model 5).

Discussion of NAEP Data Results

While NAEP reports descriptive data on student performance outcomes by looking at scores and achievement level results by gender, race/ethnicity, or SES, our analyses here explored the extent to which the variance in students' scores is attributable to demographic, SES, OTL, and students' interests and motivation toward these subjects. The greatest impact on score variance in our models came from the inclusion of demographic and SES factors. Gender, race,

and EL status together explained approximately 19 percent of the variance in students' scores. The explained variance increased by approximately 10 points when variables related to SES were added. After controlling for demographic and SES factors, adding the teacher OTL variables regarding background had relatively little impact on the explained variance of student scores, while the addition of specific classroom activities explained some additional variance in student scores, particularly in civics.

There are several possible reasons that OTL contributed relatively little to score variance in our analysis after controlling for other factors.[45] First, our measures related to teacher background (major or minor in subject area, years teaching, and recent professional development, in particular) are self-reported from questionnaires and cannot explicitly address the more substantive issue of whether teacher training programs are adequately preparing social studies teachers to teach their students the content and skills measured by the NAEP assessments. Other research has suggested that teachers themselves feel their preparation is not adequate to the task of teaching civics and civic-related practices. A recent RAND report included results from a survey that asked teachers to rate, on a scale from "not prepared at all" to "very well prepared," how equipped they felt to promote their students' civic development. [46] A little more than 60 percent of all teachers indicated they felt "somewhat prepared," and only 19 percent responded that they felt "very well prepared."[47] That same report revealed that 43 percent of secondary teachers has no preservice or in-service training on how to teach civic dispositions (attitudes related to citizenship, such as appreciation for diverse views and experiences and a recognition of the importance of activities like voting).[48] Second, the NAEP data cannot tell us where or when students gained the knowledge and skills they demonstrated on the assessments and the information captured about teacher efforts pertains only to the eighth grade. Students' understanding of these subject areas could have also been influenced by teaching practices and knowledge gained before the eighth grade.[49] Our findings suggest that while teachers are—in varying degrees—engaging in classroom practices that the research literature shows are related to learning in these domains (for example, discussing current events, engaging in debates around key topics, teaching how to evaluate different arguments related to past and current events, taking and defending positions, etc.), it is not possible to gauge through the current NAEP survey how teachers interpret or instantiate these practices in their classrooms.

The results regarding the OTL measures highlight that further work may be needed in this area so that large-scale assessments such as NAEP can capture what is happening in classrooms to help inform policy in this area. The use of self-report measures of OTL presents several challenges, including comparability of meaning across teachers who respond to these questions.[50] For example, it is not clear that teachers have a uniform way of understanding what is meant by phrases such as "using multiple sources." Second, research rightly suggests that having more versus less of some classroom activity—mock trials or debates, for example—does not necessarily translate to better learning outcomes. [51] Understanding in a more fine-grained way the relationship between frequency of exposure to specific classroom

practices and student outcomes in civics and U.S. history is a topic for further research. Such research will hopefully lead to improved teacher questionnaires in future NAEP assessments and more refined methodological approaches to analyzing NAEP survey data.

The NAEP data and analyses here suggest that (1) a large percentage of eighth grade students are not demonstrating "solid academic performance" in civics and U.S. history; (2) SES and characteristics associated with SES, such as EL status and racial/ethnic factors, are associated to learning in civics and U.S. history, as the few studies that examine these subject areas with national or state data have suggested;[52] and (3) OTL factors related to teacher background and classroom practices, at least those currently captured by the NAEP assessments, have a relatively small impact on student outcomes after controlling for demographic and SES factors. One suggestion for how to address these issues involves improved assessment design. For example, to better target interventions for lower performing students, the civics and U.S. history assessments could integrate a screening literacy test that would route students with lower levels of foundational skills (e.g., decoding, fluency) to an assessment geared toward more accurately identifying their level of knowledge and skills by presenting a more appropriate level of challenge. In addition, research regarding the relationship between the comprehension of various text- and nontext-based sources and the acquisition of knowledge and skills in these social science domains should be incorporated into the frameworks and test specifications.[53] Additional research is also warranted regarding OTL factors related to teachers' training, background, and ability to emphasize knowledge and skills related to civic learning in classrooms. Finally, we need more exploratory research on which classroom activities—and levels of using these activities—are most associated with students' outcomes and whether these vary by other student group factors (e.g., SES). Such research would hopefully inform better teacher, school, and student survey instruments and a more honed set of analytical techniques to better measure OTL in relation to civics and U.S. history.

The achievement data derived from an assessment, however, is only meaningful if it captures well the construct under consideration. What can we do to improve the data collected in both summative and formative assessments and ensure that it is measuring not only the knowledge associated with civic learning, but also the skills of critical thinking, civic participation, and civic engagement that are equally essential to creating an informed citizenry? In the next section, we explore this question in more detail.

Improving the Assessment of Skills in Civics and U.S. History

In our view, aligning assessments with core skills associated with civics and U.S. history will go a long way toward making the data collected in such assessments more meaningful and actionable. The challenge for improving the assessment of civic learning is essentially twofold: First, how do we clearly define skills and knowledge associated with the domain of civics and U.S. history; and second, how do we do so in a way that allows us to capture students' knowledge of these skills in an assessment setting. Practitioners in the fields of civics and U.S.

history education have devoted attention in the past few decades to answering the first question. This section will briefly explore this work to delineate the core skills in civics and U.S. history. The subsequent section will then examine how new item types could be used to target the measurement of these skills.

Defining Civics and History Skills

Much of the research on civics education in the past two decades has centered on a recognition of a growing sense of distrust in government and others by both adults and school-age children. This has led some scholars to warn of the dangers of a fraying social and political order. Civic education is seen as a crucial antidote to this danger. Authors of one report claim that "[C]ivic learning is an essential part of the solution" of a democracy in crisis. "In a society characterized by weak civic institutions, balkanized public discourse, and profoundly unequal civic engagement," the authors claim, "schools can offer all young people opportunities to learn fundamental facts and skills, engage with each other and with their communities, and develop dispositions and values supportive of a republican form of government."[54] Moreover, scholars and researchers note that civic education must reflect changes that have occurred since the mid-20th century in this country, including the need to (1) teach digital literacy skills, (2) engage a more diverse student body, (3) provide students with the skills they need to navigate difficult societal challenges, and (4) counter inequitable availability of resources and opportunities to learn in the K-12 level and beyond.[55]

The NAEP civics framework, while written in the early 1990s, outlines many of the skills that educators and scholars of civic learning now see as essential for students to have along with the knowledge of government. The framework refers to "three interrelated components: knowledge, intellectual and participatory skills, and civic dispositions."[56] The knowledge component of the framework covers core areas of government, including the meaning of civic life and government, fundamentals of the American government system, knowledge of constitutional democracy, and how the U.S. government interacts with other governments throughout the world. The framework defines core civic skills as the ability "to think and act effectively and in a reasoned manner in response to the challenges of life in a constitutional democracy." Skills such as the capacity to "identify, describe, explain, and analyze information and arguments, as well as evaluate, take, and defend positions on public issues" allow students to *apply* their knowledge of government in meaningful ways. According to the NAEP framework, participatory skills are those that "enable citizens to monitor and influence public and civic life by working with others, clearly articulating ideas and interests, building coalitions, seeking consensus, negotiating compromise, and managing conflict."[57] These correspond to the civic competencies defined in Figure 3, which are taken from recent research and theory in the field of civic education. Note that civic dispositions, while articulated in the various NAEP framework documents that outline core civics skills, including

the NAEP civics framework, is the skill set most challenging to measure in a standard large-scale assessment environment, although it is likely possible to glean from responses to survey question regarding students' volunteer or community-related activities.

Figure 3: Civic Competencies

Civic knowledge: understanding of government in the United States and in other nations, along with understanding of related social studies concepts, including the effects of history on current government and societies

Civics skills: ability to engage actively and effectively in democratic processes by applying skills such as critical thinking, teamwork, written and oral communication and information literacy

Civic dispositions: attitudes that support democratic participation, including an appreciation of responsibilities of citizenship, interest in the welfare of others, a sense of personal and collective agency, and capacity to engage in civil disagreement while maintaining civic friendship

Civic engagement: integration of knowledge, skills, and dispositions to solve public problems, improve communities and societies, and navigate formal and informal political systems and processes; can occur individually or collectively and encompasses civil actions and civic participation.

Source: Laura Hamilton and Ace Parsi, *Monitoring Civic Learning Opportunities and Outcomes: Lessons from an ETS/EAD Symposium* (Princeton, NJ: ETS, 2022), Figure 1, https://www.ets.org/Media/Research/pdf/Research_Notes_Hamilton_Parsi.pdf. Copyright 2022 Educational Testing Service.[58]

Historians, cognitive scientists, and others interested in history education have collaborated in last few decades to explore and theorize about the ways in which we teach and learn history. The guiding principle of this new work is that teaching and learning history is an "epistemological and cultural act."[59] In a seminal collection of essays that explored cross-nationally how history is defined, scholars in Germany, Canada, Sweden, the Netherlands, and the United States appeared surprisingly aligned on at least one central idea: The belief that the study of history at its core involved understanding *both* "content" or "historical knowledge" and a process by which individuals come to know, examine, and ultimately make sense of that content, variously termed "historical thinking," "historical consciousness," and "historical habits of mind."[60] History is understood as being at once an encounter with evidence about the past *and* an interrogation of how we construct knowledge about that past. Doing history—or learning to think historically and apply the skills essential to this domain—is part of learning history.[61]

Figure 4 highlights and briefly outlines historical thinking skills as follows: (1) the reading and analysis of historical sources, (2) skills associated with comparison and contextualization, (3) chronological reasoning to understand change over time, and (4) competencies associated with inquiry and argumentation. These four core areas are generally agreed upon by several organizations that have released documents that discuss historical thinking skills, including the Stanford History Education Group, the American Historical Association, the UCLA Public History Initiative, the College Board's Advanced Placement, and the *C3 Framework* in U.S. history, all of which are grounded in theoretical research about what it means to think historically and how to help students develop the skills they need to do so.[62]

Figure 4: Historical Thinking Skills

- **U.S. history knowledge:** understanding key themes and chronological periods of U.S. history

- **Analyzing sources and evidence:** the ability to "read" text and nontext sources, including the ability to identify, analyze, and explain various information and perspectives presented in historical sources

- **Making historical connections:** the ability to compare, contextualize, and summarize different historical contexts, time periods, sources, or perspectives of historical actors, events, and phenomenon

- **Chronological reasoning:** the ability to understand, analyze, and interpret issues related to causation and how phenomena change over time and in relation to events and circumstances

- **Creating and supporting a historical argument:** the ability to interpret and evaluate historical arguments and develop interpretations of arguments that are supported with evidence.

Note: Adapted from College Board, *AP Historic Thinking Skills* (New York: College Board, 2015) https://www.acpsd.net/site/handlers/filedownload.ashx?moduleinstanceid=6536&dataid=25384&File Name=historical_thinking_skills_1_.pdf, "Historical Thinking Skills," American Historical Association, accessed January 23, 2023, https://www.historians.org/teaching-and-learning/teaching-resources-for-histo rians/teaching-and-learning-in-the-digital-age/the-history-of-the-americas/the-conquest-of-mexico/for- teachers/setting-up-the-project/historical-thinking-skills; National Council for the Social Studies, *The College, Career, & Civic Life (C3) Framework for Social Studies State Standards: Guidance for Enhancing the Rigor of K-12 Civics, Economics, Geography, and History* (Silver Spring, MD: National Council for the Social Studies, 2013), https://www.socialstudies.org/sites/default/files/c3/C3-Framework-for-Social-Studies.pdf.

Developing New Interactive Tasks and Tools

Although traditional question types such as multiple choice/selected response and constructed response can be designed to measure some of the skills described in the previous section, the complexity of these skills will likely require the development of new types of assessment tasks in civics and U.S. history domains. Borrowing from advancements in measurement science and question development work in other assessments, we explored the merits of designing new kinds of tasks and tools that might enhance the measurement of cognitive skills and reasoning processes civics and U.S. history.[63] Within such tasks, students

can engage with domain-relevant materials in a meaningfully structured way and answer a sequence of items that enable them to demonstrate various knowledge, skills, and abilities; use strategies and tools of the discipline; and apply, synthesize, or extrapolate from evidence or information acquired in the task to solve a problem, answer questions, or take action. By using simulated versions of authentic, disciplinary contexts, these types of assessment tasks can potentially provide more precise evidence of what it looks like when students engage in the work of the discipline.[64] For example, students might work with virtual peers or experts to come to a consensus about a proposed bill and present their recommendations to a town council (e.g., whether members should vote for or against the bill). They might first answer some standard selected-response and constructed-response questions early in a task and then apply information, knowledge, and skills to complete additional work products, such as writing an editorial, developing a plan to resolve a community conflict, or gathering evidence to support an argument about a civics topic or historical interpretation. The combination of multiple item formats allows each task to be purposefully designed to measure multiple skills across a range of difficulties and complexity levels. In addition, evidence suggests that such interactive, dynamic tasks can be more engaging to students than traditional assessments, with greater engagement expected to yield better, more precise evidence of what students know and can do.[65]

Digital tools also can simulate disciplinary contexts that range from the simple to the complex and can be used in standalone items, in sets of items, or within a larger task design. For example, students can interact with multimedia sources, dynamic data displays, or simulated web search tools in the service of solving disciplinary problems. Another benefit of digital tools is that they can produce rich process data—digital records of the sequence and timing of all test-taker actions—that can yield evidence of students' moment-by-moment thinking and reasoning as they unfold in the course of completing the task.[66] Tasks using digital tools can be designed with process data in mind and can thus provide additional insights into students' skills and possible problem-solving strategies. This information can then be leveraged to improve pedagogy and future assessments. Below are some examples of the new interactive tasks that could be developed to improve the measurement of core skills in civics and U.S. history. The tasks were written specifically for this report to illustrate how tasks—and tools embedded within tasks—can be used to measure the critical thinking skills in U.S. history and civics in grades 8–12 that are outlined in the previous section.

Exploring Multiple Sources

Several large-scale assessments in a variety of subjects have utilized a tabbed interface to allow students to explore, analyze, and evaluate multiple sources of information pertaining to a specific topic.[67] We envision the use of a "source container" tool, the primary feature of which would be to make multiple text and nontext sources available within a task by placing them in easily accessible tabs among which students can toggle. The sources could be of any type (text, images, audio, video, data, etc.). The introduction of digital sources would help

extend the representation of target constructs to include measurement of a broad range of digital and media literacy skills. Depending on the sources selected and the questions asked, virtually all the skills referenced as central to historical thinking in Figure 4 and many participatory and engagement skills in civics outlined in Figure 3 can be assessed using this type of tool. In history, where integrating multiple sources is at the heart of constructing historical narratives and analyses, this type of tool could facilitate the measurement of skills ranging from basic understanding of source material to chronological reasoning, perspective-taking, and critical evaluation and argumentation. In civics, this type of tool could be useful in targeting civics interacting skills like building knowledge and deliberating on public issues (e.g., by weighing evidence presented across multiple sources representing different perspectives) and monitoring skills like evaluating what the government is doing (e.g., by reading multiple sources that discuss the actions of the government or other stakeholders to track progress on an issue or a bill).

Assessment tasks can be introduced with scenarios or more simply as a series of items around a common theme or topic. Depending on the measurement goals, the sources can be introduced one at a time in a sequential presentation or all made accessible at the beginning of the task. Each source can be accompanied by background information or metadata (e.g., author or creator, date of publication or airing, purpose) that is useful in contextualizing the source and determining its trustworthiness. In tasks where sources are introduced sequentially, the background information can first appear on a preview screen and then remain available for the remainder of the task via a pop-up window accessible by pressing an "information" button. Data on whether students access the background information at other points in the task can be collected—either from process logs or questions that focus specifically on the use of background information—to help probe students' ability to place documents in historical context and connect historical developments or themes to a specific time and place. All standard item types used in large-scale assessments, including interactive ones, would be compatible with the source con. The source container envisioned here could expand the functionality of the select-in-passage item type that has been developed for NAEP reading assessments[68]—in which students can directly select portions of a text-based source to serve as evidence in support an argument or point of view—to include a variant of the select-in-passage item type that is compatible with video and audio sources as well as multiple tabs. As a result, students can be presented with a given claim or argument about a film or audio clip, or asked to make an argument or draw a conclusion of their own, and can then be asked to choose the portion(s) of video or audio sources that serve as supporting evidence for the argument, critical thinking skills that essential components in both history and civics.

For the purposes of demonstrating the kinds of knowledge and skills that a source container tool can allow us to measure, we now describe an example of a U.S. history task developed for this paper on the topic of Freedom Summer, the civil rights effort to increase African American

voter registration in Mississippi in the summer of 1964, during which three volunteers were kidnapped and murdered. As shown in Figure 5, students taking the assessment could receive basic contextual information on the first screen of the task.

Figure 5: Source Container Example: Introductory Screen

In the summer of 1964, the Council of Federated Organizations (COFO), a coalition of the Mississippi branch of the four major civil rights groups, organized the Mississippi Summer Project. Freedom Summer (as it was later known) was a volunteer project, the goal of which was to register as many African American voters as possible in Mississippi, where African Americans had been restricted from voting since the turn of the century. Bob Moses, a civil rights leader and co-director of COFO, led the project.

For this activity, you will examine several historical sources about the Mississippi Summer Project (1964).

Note: The material in the figure does not represent actual content developed for NAEP or any other assessment.

Students would then be introduced to three historical sources, as shown in Figure 6: (1) a video of a CBS news program on Freedom Summer that aired after the initial disappearance of the three civil rights volunteers in Philadelphia, Mississippi; (2) an audio recording of interviews conducted with Mississippians by one of the Freedom Summer volunteers in 1964; and (3) a letter that Bob Moses, director of the Student Non-Violent Coordinating Committee, sent to parents after the volunteers had disappeared. The sources would be presented one at a time, with each source preceded by the background information. Figure 6 shows the film information screen preceding the introduction of the video.

Figure 6: Source Container Example: Film Information Screen

Note: The material in the figure does not represent actual content developed for NAEP or any other assessment.

In this task, the film clip would be accompanied by two questions. Figure 7 shows the first question in on the right side of the screen and the video player on the left, in a tab. Pressing the film info button in the video player opens a pop-up window containing the film information from the previous screen. The purpose of the first question would be to obtain a baseline for whether students can read and comprehend some of the basic information in the source. This aligns to the historical thinking skill of understanding and being able to identify useful information from historical sources.

Figure 7: Source Container Example: Multiple-Choice Question Measuring Understanding of Film Clip

Note: The material in the figure does not represent actual content developed for NAEP or any other assessment.

The second question, shown in Figures 8 and 9, elicits evidence of a deeper understanding of the source and the demonstration of historical thinking skills related to source evaluation and historical argumentation. The question is structured as a composite item. The first part is a multiple-choice item asking students to identify an argument made in the news clip. The second part is a select-in-video item in which students must select the portion of the news clip containing evidence supporting the argument they chose in the first part. This proposed item type is built on a "select-in-text" functionality that some large-scale assessments have used.[69] In Figure 8, one can see the film clip divided into four segments. Students would be able to rewatch the entire video or any of the individual segments before answering questions related to this source. When they choose a segment as their answer, it would appear in the answer space on the right side of the screen. In this item, there is a single correct answer to each part, but these types of questions may intentionally have more than one correct answer because the purpose is to elicit information about students' ability to support their opinion, position, or argument with appropriate evidence. This design feature would enhance the measurement of certain skills related to history and civic learning by acknowledging that there can often be more than one correct answer, and the correctness of a response can depend to the ability of students to support their opinion or view with corroborating evidence and valid reasoning. These kinds of composite items are useful in that they can align to multiple skills, such as the ability to understand or describe the argument of others, or provide evidence to support a proposed argument, which are key critical thinking skills embedding in US history and civics.

Figure 8: Source Container Example: First Portion of Composite Item, Showing Multiple-Choice Question

Note: The material in the figure does not represent actual content developed for NAEP or any other assessment.

Figure 9: Source Container Example: Second Portion of Composite Item, Showing Video Clip Selected as Evidence
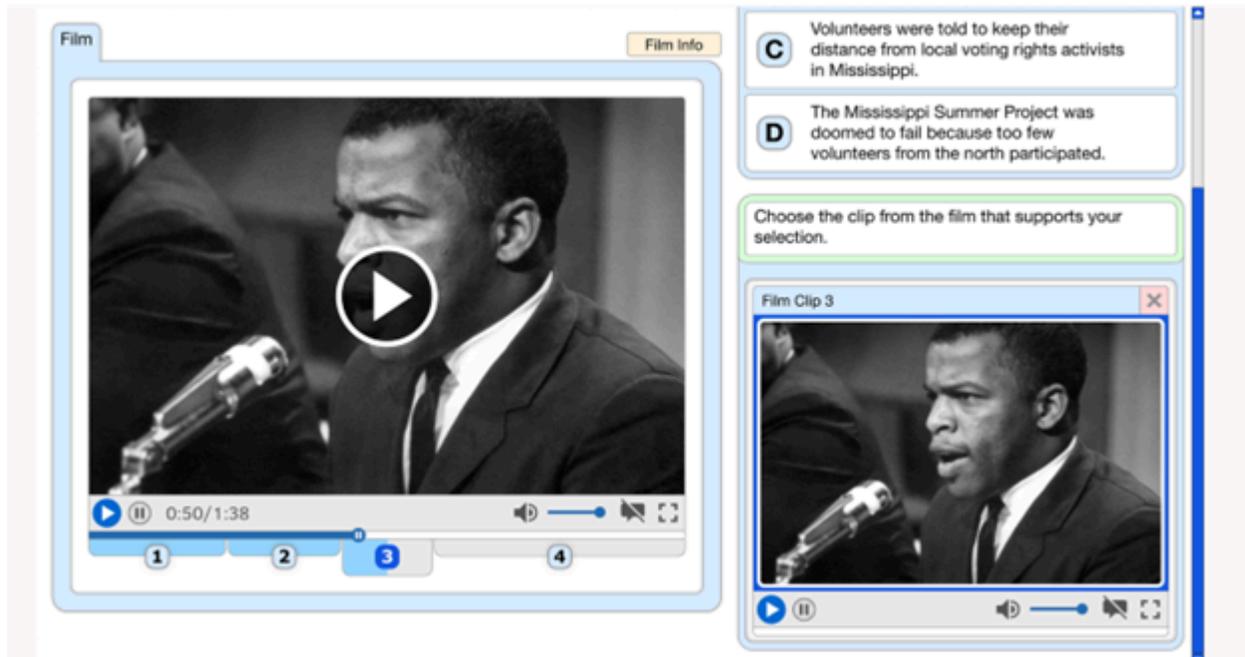
Note: The material in the figure does not represent actual content developed for NAEP or any other assessment.

After answering the two questions related to the film clip, students would be introduced to the radio broadcast and asked to answer one or two questions about it or possibly about it and the film clip. Then the final source, the letter from Bob Moses to parents of the Freedom Summer volunteers, would be introduced. This would also be followed by one or two questions focused on that document, perhaps asking why Moses had written the letter or why he believed federal protection was needed for the volunteers. The task would then pose some final questions for which students would need to consult and compare all three sources. One example is displayed in Figure 10. Like the question displayed in Figures 8 and 9, this item involves having students find and select relevant evidence in support of a given claim. Though not illustrated here, this item could be followed by a final constructed-response item asking students to develop their own argument supported by evidence or to engage in some other activity requiring a degree of synthesis of the three sources and accompanying contextual information, which would enable measurement of different aspects of historical thinking or civic learning skills.

Figure 10: Source Container Example: Question Measuring Ability to Support a Claim with Evidence and Allowing the Use of All Sources

Note: The material in the figure does not represent actual content developed for NAEP or any other assessment.

Simulated Search

A simulated search tool, discussed below, could be used to measure interacting and monitoring skills like researching public issues, gathering and analyzing information, and using electronic resources. Given the ubiquitous nature of digital information, it has become clear that a core aspect of civic literacy is being able to evaluate whether students can find and select relevant, useful, and trustworthy information on the Internet—in other words, skills related to conducting research and inquiry into civic issues within online environments.[70] These are skills that students of all ages appear to find challenging, but which are essential for participation in contemporary civic life.[71]

A simulated search tool could provide a simulated search engine for conducting searches and a simulated web browser for viewing websites. The simulated search engine would allow test takers to enter search terms in natural language and receive a set of search results or hits related to the terms used, similar to a typical Google search in appearance but far more constrained in its actual implementation in order to control the presentation of information and to support measurement goals.[72] The search results would include website titles, brief descriptions of the website contents, and URLs (which for security reasons would not link to actual live websites). Test takers could then select the most appropriate website(s) for the assigned task. The search engine tool could be used on its own, so that the entire task is focused on the skill of conducting searches to gather information (i.e., constructing relevant search terms, selecting relevant websites from the search results) or it could be combined

with a simulated web browser component, which would present students with the content of one of the websites they were searching for or selecting in the search engine. Websites presented within the simulated web browser tool can have graphics, hyperlinks, and subpages, much like typical websites, and students would be able to freely navigate among these different elements as they read and explore the websites. Using this part of the tool, students could be presented with a series of questions targeting skills such as navigating among the site's pages to locate, gather, and analyze information; understanding and assessing the positions and viewpoints of others; and developing and defending their own positions, using evidence to support them. In other words, a wider range of interacting and monitoring skills, central to civic learning, could be measured by creating tasks that use both parts of the simulated web search tool (i.e., searches and websites). There can be a good deal of flexibility in how many questions to include in the second part of the task depending on the topic, richness of the source material, available assessment time, and measurement goals.

To illustrate the key features of such simulated web search tools, we will describe an example civics task about whether a school district should lengthen its school year (and shorten its summer vacation). In Figure 11, students are informed about the potential lengthening of their school year and the need to learn more to enable them to become involved in the public conversation on the issue. They are then instructed to use the simulated web search tool to search for and select the most relevant website and to try to do so efficiently. In Figure 11, the search term "summer vacation" is relevant but too general on its own to yield the desired websites. By adding the relevant keywords "schools" and "shorten," as shown in Figure 12, the search terms successfully yield some relevant results.

The search portion of this task we are envisioning could generate two separate scores. The first score would be derived from the number of relevant search terms used (since one or two-word searches are too imprecise to yield the desired websites) as well as the number of searches performed. The scoring rubric could be written to give partial credit for searches that include some number of relevant keywords short of the full complement needed to return the desired hits (i.e., the search "school shorten summer vacation" would be given full credit and the search "school vacation" might be awarded partial credit). The second score could reflect whether the student selects the correct (i.e., most relevant) website from among the search results. Depending on the topic of the task and how it is constructed, the approach to scoring can be flexible so that, for example, there is more than one correct website from which to choose.

Figure 11: Simulated Search Example: Introductory Screen
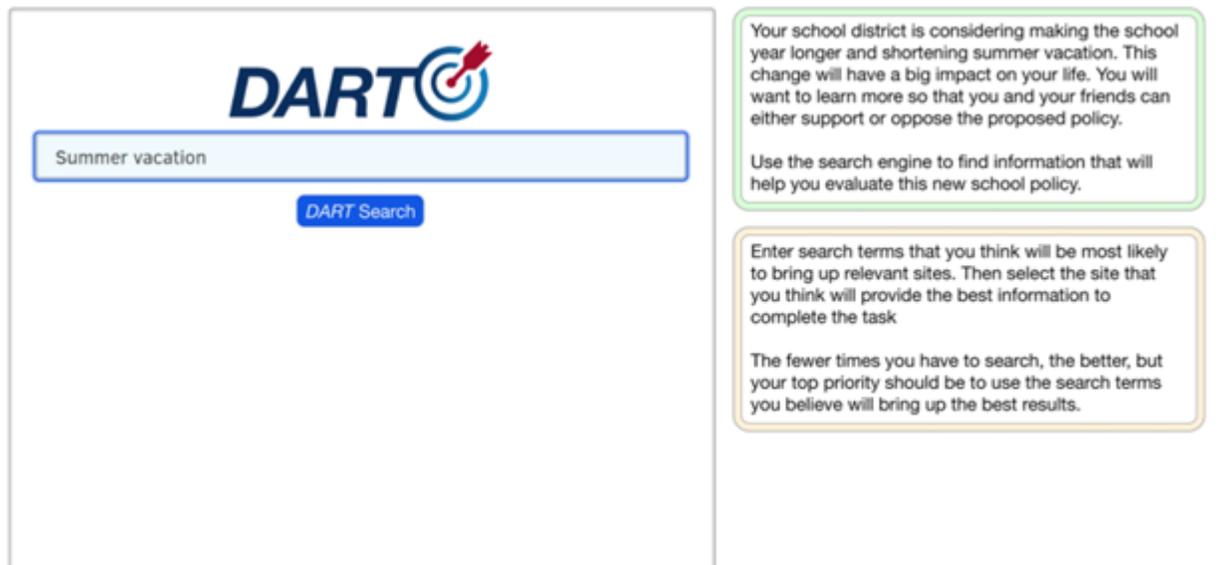
Note: The material in the figure does not represent actual content developed for NAEP or any other assessment.

Figure 12: Simulated Search Example: Screen Showing Search Results
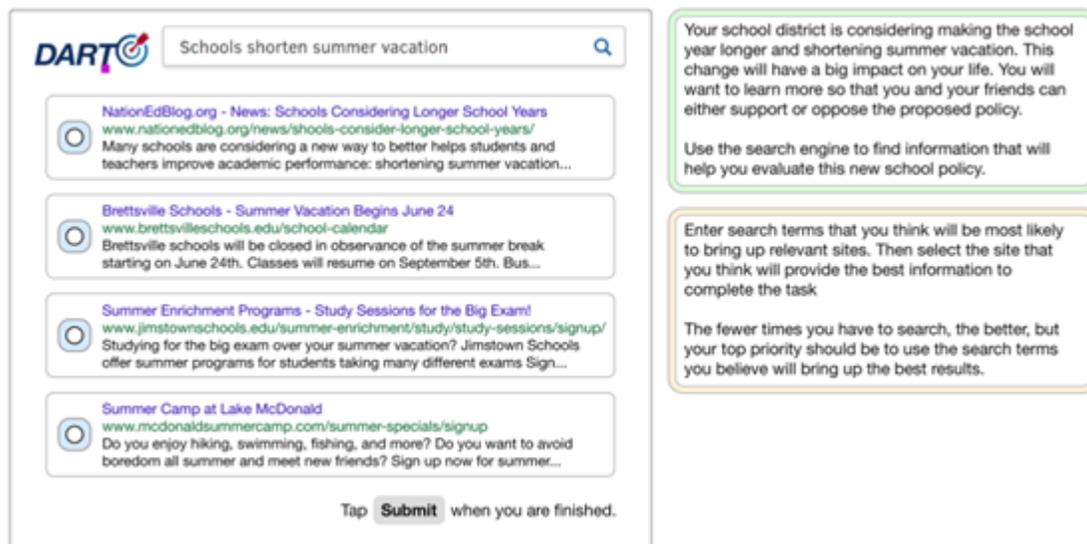
Note: The material in the figure does not represent actual content developed for NAEP or any other assessment.

After making their selection in the search portion of the task and tapping the submit button, all students could be leveled or "reset" as they move into the second part of the task with the simulated web browser. That is, regardless of how students responded to the previous web search item, they would all be presented with the same website to explore, to ensure that all students view the same materials going forward. Students would also then not be able to go back to the search portion of the task. A transition screen would briefly orient students to the

next section of the task. As part of the leveling process, students could be told that a friend found the website, thereby avoiding the need for students to think about whether they had selected the correct website in the search section.

Figure 13 shows the landing page for the web browser part of the task, with the simulated web browser tool and a website on the left. Students would be given instructions on navigating within the website, which contains multiple pages and hyperlinks, and are then asked a series of questions requiring them to locate, analyze, and evaluate the information contained on the website.

Figure 13: Simulated Search Example: Web Browser and Navigation Tips

Note: The material in the figure does not represent actual content developed for NAEP or any other assessment.

The first question, shown in Figure 14, is designed to have students become familiar with the website and topic involved in this task. It asks students to identify a primary reason offered on the website in support of a longer school year; thus, it measures whether students can extract relevant information and understand the arguments and conclusions of media sources, skills that are essential to the civics engagement. Although the figure shows a video in the web browser, students would be expected to draw from information provided in all the tabs on the website to answer the question.

Figure 14: Simulated Search Example: Multiple-Choice Question Measuring Basic Understanding

Note: The material in the figure does not represent actual content developed for NAEP or any other assessment.

The second question in this part of the sample task is a constructed-response question asking students to explain which speaker in the video they find most convincing (Figure 15). It is meant to measure how well students can evaluate both the quality of the speakers' arguments and their trustworthiness, skills associated with civic participation. Test takers can choose any of the people interviewed but their responses would be scored on the quality of the explanation they provide.

Figure 15: Simulated Search Example: Constructed-Response Question Measuring Ability to Evaluate Arguments

Note: The material in the figure does not represent actual content developed for NAEP or any other assessment.

The final question, also open-ended, connects back to the task scenario and measures students' ability to construct a convincing, well-supported argument that might influence the school district's decision (see Figure 16). Students can take any position on the issue, but their responses would be evaluated based on how well they present their argument and use supporting evidence from the website. This item aligns to the influencing skill category, more specifically "communicating information, ideas, and arguments in order to affect politics, governance, and community issues."

Figure 16: Simulated Search Example: Constructed-Response Question Measuring the Ability to Take and Support a Position

Note: The material in the figure does not represent actual content developed for NAEP or any other assessment.

Timeline Map

A digital timeline-map tool could also be used to measure a number of key historical thinking skills, including chronological reasoning, historical causation, continuity and change, and contextualization and comparison. This type of tool uses a series of maps to dynamically show changes in a phenomenon over time and space.[73] A slider mechanism could be developed to allow the user to advance through the years in order to visualize the changes. Maps can be drawn to various temporal (century, decade, year, month, etc.) and geographic scales (world, nation, state, community, etc.) depending on the phenomenon being depicted. A simple version of the tool could rely on a series of static maps, but more sophisticated versions can draw from a database, allowing for smoother transitions between data points and zooming in to explore areas in greater detail. Questions based on the timeline-map tool can address larger patterns of change over the entire time span represented or focus on more local patterns and events and shorter periods. Guided by the measurement goals, questions can direct students to look at specific maps or points in time or leave it up to students to identify which maps they need to consult to answer a question. Additional stimulus material can also be added to individual items, to be analyzed along with the maps, in support of measuring students' contextualization or comparison skills.

To illustrate how the timeline-map tool could be incorporated into tasks that measure historical thinking skills, we next describe an example U.S. history task about railroad development from 1830 to 1890 and its relationship to urban population growth. In this

sample task, maps are displayed at 10-year intervals, with seven distinct maps available to be displayed. Three of many possible questions are given as examples. The first question (see Figure 17) asks students to draw a conclusion about the nature of changes in the growth of both railroads and cities between 1830 and 1860, a task inspired by a released grade 8 task for the NAEP TEL assessment.[74] It is designed to probe students' understanding of the nature of historical causation, the limits of certain source materials, and ability to engage in chronological and reasoning, skills that can be usefully probed using interactive maps.[75]

Figure 17: Timeline-Map Example: Interface and Multiple-Choice Question
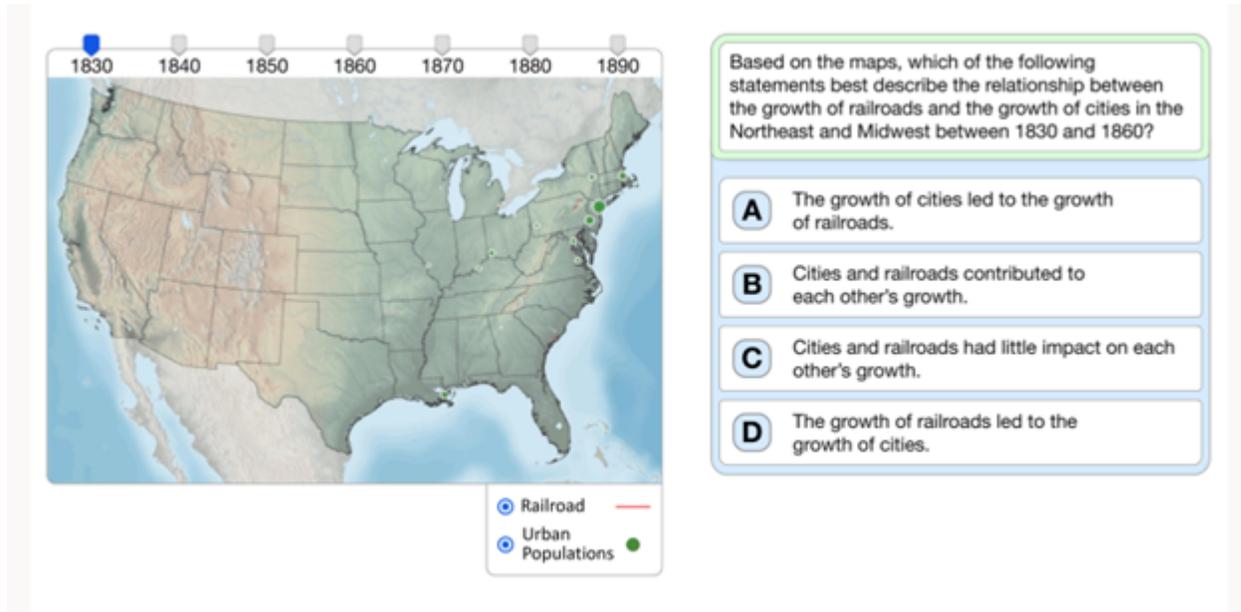
Note: The material in the figure does not represent actual content developed for NAEP or any other assessment.

The second question (Figure 18) asks about the importance of Chicago as a freight and passenger hub by the year 1870. To provide a successful answer to the question, students would need to use the maps to evaluate the geographic pattern of railroad development and population growth and bring to bear historical knowledge of migration, trade, and commerce, as well as the ability to analyze and evaluate cause and effect relationships.

Figure 18: Timeline-Map Example: Composite Constructed-Response Question Measuring Map Reading and Historical Contextualization
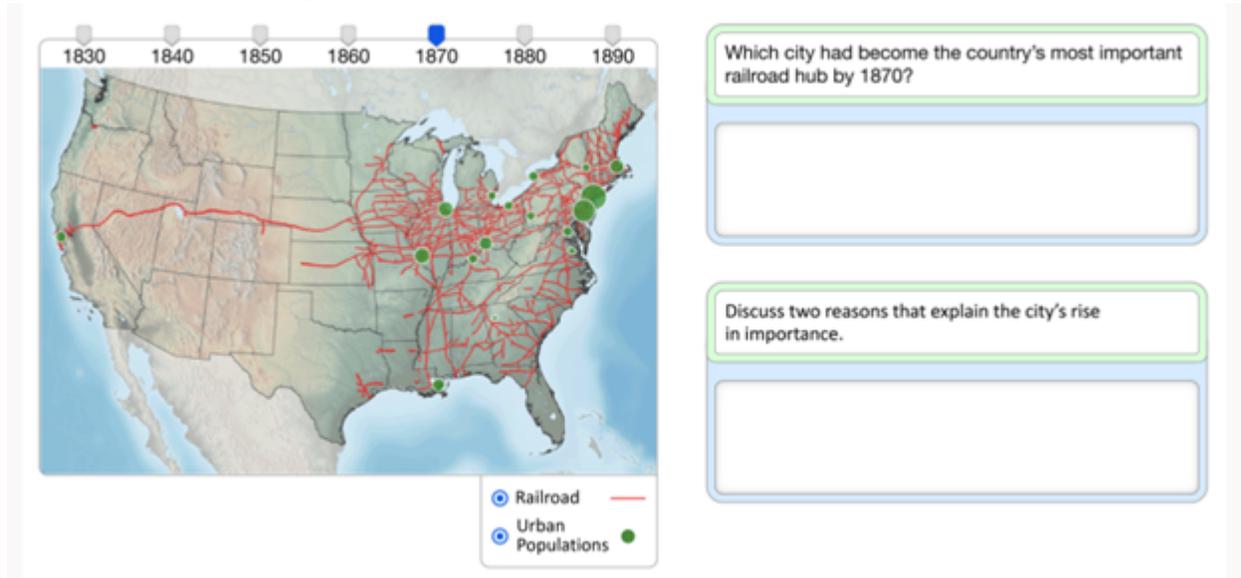
Note: The material in the figure does not represent actual content developed for NAEP or any other assessment.

Students are commonly taught that a relative lack of railroad development in the South versus the North was a disadvantage to the Confederacy during the Civil War. In the final sample question (Figure 19), students are essentially asked to extend this analysis out to 1870. Although the question refers specifically to the South, students in fact need to compare the lack of development there to continued development in the North and to contextualize Southern stasis within the events of the 1860s (e.g., the physical devastation from the war, a reluctance among influential segments of the population to transition the economy away from agriculture, and the displacement of large portions of the workforce). The question thereby combines the measurement of students' knowledge of the Civil War and Reconstruction era with their ability to connect historical developments to specific circumstances in time and place, and to broader regional or national processes. It aligns to the important skill of understanding historical causation and patterns of continuity and change.

Figure 19: Timeline-Map Example: Composite Constructed-Response Question Measuring Historical Knowledge and Chronological and Causal Reasoning
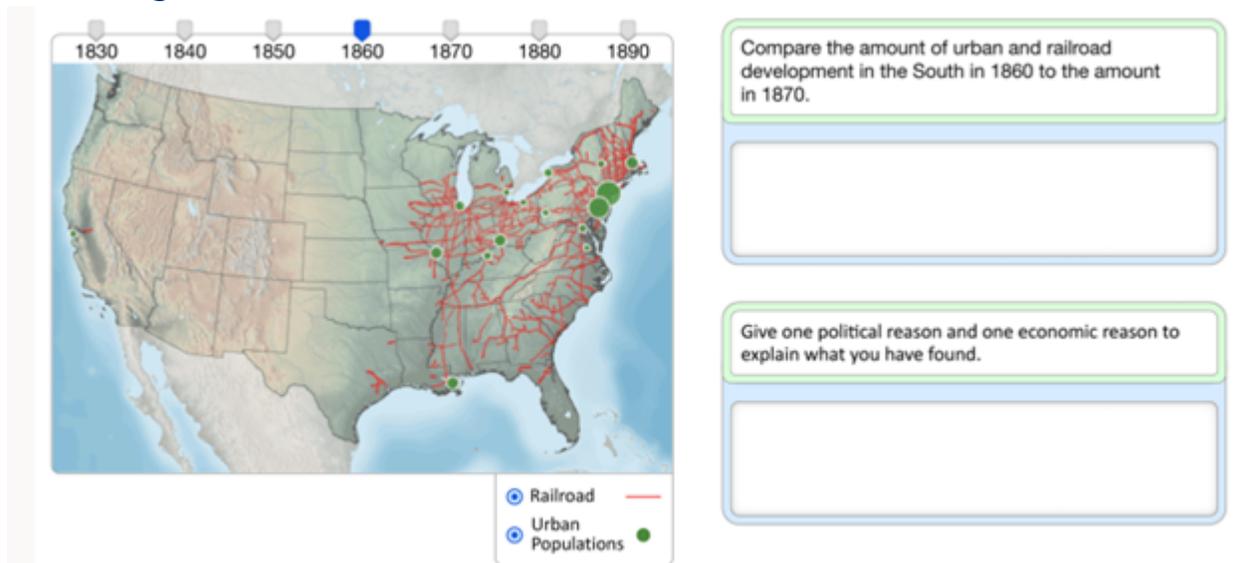
Note: The material in the figure does not represent actual content developed for NAEP or any other assessment.

Summary and Implications

It is likely no accident that the redefinition of American citizenship embodied in the Fourteenth Amendment to the Constitution, passed by the Senate in 1866 and ratified in 1868, occurred at the same time as the establishment of the federal Department of Education in March of 1867. The Department was designed "for the purpose of collecting such statistics and facts as shall show the condition and progress of education in the several States and Territories, and of diffusing such information respecting the organization and management of school and school systems, and methods of teaching, as shall aid the people of the United States in the establishment and maintenance of efficient school systems, and otherwise promote the cause of education throughout the country."[76] It took a century before the Department of Education fully acted upon this purpose, particularly in the area of collecting data to addresses "the progress of education."[77] In 1969, the first trial national assessments were given in citizenship, science, and writing. At core, then, education in citizenship and other disciplines—and the sound, reliable measurement of these subjects throughout the nation—are deeply rooted together. This paper investigated how we can utilize the data that NAEP civics and U.S. history assessments currently provide to better understand the relationship between performance and other key in-school and out-of-school factors. In addition, we have examined how we can improve the large scale and summative measurement of many of the skills central to the domains of U.S. history and civics. We believe

doing so will support recent efforts to redefine how civic learning is taught and encourage the development of critical thinking skills in our students—skills that are necessary to support democratic citizenship in the future.

Our analysis of the 2018 results in civics and U.S. history reveals several key points. Descriptive data from the assessments show that while there have been improvements for some subgroups of students over time, achievement gaps remain among ethnic/racial groups and those of differing socioeconomic status. Moreover, there are large percentages of students overall who struggle to demonstrate higher order skills such as critical reading, interpretation, and argumentation. The regression analyses, which accounted for key demographic, SES, OTL, and student attitudes related to core skills in civic learning, revealed that nearly 40 percent of the variance in student scores was attributable to these factors (39 percent in NAEP civics and 36 percent in NAEP U.S. history).The analyses also underscored that opportunity broadly defined—and educational opportunity in particular—is unequally distributed across demographic and SES groups in the United States. The demographic and SES factors alone accounted for approximately 30 percent of the overall variance in students' scores. Although a small percentage of the overall variance could be attributed to the OTL factors used in the analyses after demographic and SES variables were taken into account (4 percentage points for civics and 1 percentage point for U.S. history), other research has shown that OTL is intricately bound up with demographics and SES. Schools in economically disadvantaged urban and rural districts, for example, often struggle to hire and retain quality teachers, have access to fewer resources, and face other challenges that can infringe on students' opportunities to learn.[78] This more complicated view of OTL suggests the possible need for improvement in NAEP's OTL measures and approaches to the analysis of data derived from the OTL-related survey questions. Above and beyond specific implications for the improvement of teacher-based OTL measures, we must acknowledge that to truly have an impact on students' civic learning and to address growing inequality in education we will need to make reforms both inside and outside of the school doors.

The second part of the paper focuses on aligning civics and U.S. history assessments more closely with the skills that experts and practitioners in the field believe are most critical in these domains. The discussion of how we can better measure civic participatory and engagement skills and historical thinking skills *along with* knowledge of civics and U.S. history content highlights challenges for large-scale assessment and measurement of these constructs overall. The usefulness of the data that any assessment provides can only truly be evaluated in light of how the domain of knowledge is conceptualized in the construct definition (i.e., in the frameworks) and instantiated in assessment items. Important recent initiatives in civic learning highlight the need to emphasize higher order domain skills, along with civic and U.S. history knowledge, but may require more support and specificity regarding their integration into instructional programs and assessments. Both the *C3 Framework* and aspects of the *Roadmap for American Democracy*, two key documents in the field, stress the necessity of teaching higher order skills. The *Roadmap* defines a number of key "themes" and

"design challenges" that incorporate skills. For example, Theme 1 addresses civics participatory skills, and Theme 7 addresses how "historical narratives shape current political arguments, how values and information shape policy arguments, and how the American people continues to renew or remake itself in pursuing fulfillment of the promise of constitutional democracy."[79] This understanding of how historical narratives function involves many of the components of historical thinking skills described previously, including chronological reasoning, evaluating evidence, and making and supporting arguments about change over time. The *Roadmap*'s Design Challenge 2: America's Plural yet Shared Story integrates the skill of analyzing multiple perspectives that is articulated in numerous places in the historical thinking skills model. Finally, Design Challenge 5: Balancing the Concrete and the Abstract touches on the complex relationship between knowledge and more abstract application of skills in civics learning overall.

The current attention given to issues of citizenship and education harkens back to the period of the passage of the Fourteenth Amendment in the years following the Civil War, when citizenship and education were inextricably bound together in explicit ways. As constitutional scholar Derek Black has argued, "The original intent behind the Fourteenth Amendment included a commitment to guarantee education as a core aspect of state citizenship...[and] that education was just as important to securing full citizenship as voting. Simply put, the Fourteenth Amendment guaranteed citizenship, and citizenship required education."[80] The right to citizenship is therefore in an important sense the right to quality education that allows for the full expression of that citizenship. This requires that citizens have knowledge of our government and history and the skills needed to examine, interrogate, and synthesize competing information as well as the ability to act responsibly as members of society. As a nation, we need to commit ourselves to providing high quality civics and U.S. history education for all students. Monitoring student progress with assessments that measure the knowledge and skills students need to understand and actively engage in the protection of their rights and an articulation of their responsibilities is a key part of this commitment to promote a healthy democracy.

Appendix A: Achievement Level Descriptors for NAEP Grade 8 Civics

LEVEL	DESCRIPTION
NAEP Basic **(134)**	Eighth grade students performing at the NAEP Basic level should have some understanding of competing ideas about purposes of government, and they should be able to describe advantages of limited government. They should be able to define government, constitution, the rule of law, and politics. They should be able to identify the fundamental principles of American democracy and the documents from which they originate, and they should understand the importance of a shared commitment to the core values of American democracy. They should recognize the components of the political process and understand personal, political, and economic rights and responsibilities. They should be able to describe the purposes of some international organizations.
NAEP Proficient **(178)**	Eighth grade students performing at the NAEP Proficient level should understand and be able to explain purposes that government should serve. These students should have a good understanding of differences between government and civil society and of the importance of the rule of law. They should recognize discrepancies between American ideals and reality and be able to describe continuing efforts to address them. They should understand the separation and sharing of powers among branches of government and between federal and state governments, and they should be able to explain how citizens influence government. They should be able to describe events within the United States and other countries that have international consequences.
NAEP Advanced **(213)**	Eighth grade students performing at the NAEP Advanced level should have a developed understanding of how civil society helps to maintain limited government and why the rule of law is important. These students should have a clear understanding of issues in which democratic values are in conflict and of past efforts to address the discrepancies between American ideals and reality. They should understand how citizens can monitor and influence government and how responsible citizens support democracy. They should recognize the impact of American democracy on other countries, as well as other countries' impact on American politics and society.

Appendix B: Achievement Level Descriptors for NAEP Grade 8 U.S. History

LEVEL	DESCRIPTION
NAEP Basic **(252)**	Eighth grade students performing at the NAEP Basic level should be able to identify and place in context a range of historical people, places, events, ideas, and documents. They should be able to distinguish between primary and secondary sources. They should have a beginning understanding of the diversity of the American people and the ways in which people from a wide variety of national and cultural heritages have become part of a single nation. Eighth grade students at the NAEP Basic level should also have a beginning understanding of the fundamental political ideas and institutions of American life and their historical origins. They should be able to explain the significance of some major historical events.
NAEP Proficient **(294)**	Eighth grade students performing at the NAEP Proficient level should be able to explain the significance of people, places, events, ideas, and documents, and to recognize the connection between people and events within historical contexts. They should understand and be able to explain the opportunities, perspectives, and challenges associated with a diverse cultural population. They should incorporate geographic, technological, and other considerations in their understanding of events and should have knowledge of significant political ideas and institutions. They should be able to communicate ideas about historical themes while citing evidence from primary and secondary sources to support their conclusions.
NAEP Advanced **(327)**	Eighth grade students performing at the NAEP Advanced level should recognize significant themes and movements in history and begin to understand particular events in light of these themes and movements. They should have an awareness of continuity and change over time and be able to draw relevant analogies between past events and present day situations. They should be able to frame questions about historical topics and use multiple sources to develop historical generalizations and interpretations. They should be able to explain the importance of historical themes, including some awareness of their political, social, and economic dimensions.

Appendix C: Methodology

The data used in this analysis came from the 2018 8th grade civics and U.S. history assessments from NAEP. The assessments include multiple-choice, short-response, and constructed-response subject-specific cognitive items as well as a student background questionnaire with general and subject-specific contextual questions. Students received different blocks of cognitive items through a technique called balanced incomplete block (BIB) spiraling in which all items are completed by a sample of students, though individual students only complete a fraction of the items. Nationally representative samples of students take various portions of the entire pool of assessment items using matrix sampling. Both student performance data and questionnaire responses were used in this analysis to gain a more nuanced understanding of the NAEP civics and U.S. history results and their relationship to demographics, opportunity to learn, and select student attributes.

The assessment was implemented through both digitally based and paper-based administrations. In 2018, approximately 13,400 eighth graders from 780 schools took the civics assessment and 16,400 eighth graders from 780 schools took the U.S. history assessment. The nationally representative sample was generated through multistage cluster sample designs to allow for accurate estimates of student populations.[81] However, due to missing data rates for certain questionnaire items from paper-based administrations, only data from digitally based administrations were used in this analysis. The sample sizes were 6,164 students for the final U.S. history regression model and 3,420 students for the final civics regression model.

For student performance data, the NAEP metric used for this analysis was average scale scores, which can range from either 0–500 (U.S. history) or 0–300 (civics). These scales were produced using item response theory methods, which can produce plausible values that are approximate estimates of students' overall performance. A composite of five plausible values was used per subject to create a single measure of student performance in this analysis.

NAEP questionnaire items were used to contextualize student performance through indicators of demographics, OTL, and student attributes. The NAEP questionnaire items utilized in this analysis came from the U.S. history student questionnaire, civics student questionnaire, social studies teacher questionnaire, and social studies school questionnaire.[82] A total of 18 questionnaire items were selected for U.S. history, and 21 items for civics. Variables were divided into five categories that included demographic characteristics, indicators of socioeconomic status, a cluster of opportunity to learn variables related to teacher background and teacher classroom emphasis on particular skills associated with civics learning, and selected student indices that reflect students' interest, motivation, and confidence in civics and U.S. history.

We ran preliminary descriptive and correlational analyses to estimate the relationship between scale scores and variables associated with SES, OTL, and student attitudes related to core skills in civics and U.S. history. We then selected variables for our regression models.

Once key variables were selected, hierarchical multiple linear regression was used to provide greater information on the relationship between variables, identify and isolate the impact of each category of variables, and control for other influences. To ensure regression was an appropriate approach, assumptions of normality, linearity, independence, homoscedasticity, and multicollinearity were investigated, and all assumptions were met.

The statistical models were estimated using AM, a statistical software specifically designed for analyzing large-scale assessment data,[83] which can correctly apply sampling weights and plausible values for complex samples. All statistical models incorporated 2018 NAEP-created sampling weights. To compare between nested regression models, an F-test was used to determine if the change in R^2, or the amount of variance explained in the dependent variable, was significant (see Table 2 for these results).

The regression analyses investigated the relationship between student performance (measured as a composite of five scale score plausible values) and five categories of variables: (1) demographics, (2) SES indicators, (3) teacher background/preparation, (4) teaching emphasis, and (5) student intra/interpersonal factors. Categories 3 and 4 include variables related to opportunity to learn. The specific variables found in each category are described in detail in Tables C2 and C3. Each category was treated as a block of variables and entered hierarchically to produce five regression models, such that each model included an increasing number of blocks (see Table C1 for a list of regression models per subject). This was repeated for each subject, where five models were tested for civics and five models for U.S. history. See Tables C4 and C5 for detailed findings.

Appendix Table C1: Variable Clusters Included in Each Regression Model

MODEL 1	MODEL 2	MODEL 3	MODEL 4	MODEL 5
Demographics	Demographics	Demographics	Demographics	Demographics
—	SES indicators	SES indicators	SES indicators	SES indicators
—	—	Teacher background/ preparation	Teacher background/ preparation	Teacher background/ preparation
—	—	—	Teaching emphasis	Teaching emphasis
—	—	—	—	Student intra/inter-personal factors

NOTE: SES= socioeconomic status.

Appendix Table C2: Civics Independent Variables List

VARIABLE (NAEP ID)	QUESTIONNAIRE	ITEM OR INDEX	REGRESSION CATEGORIES
DEMOGRAPHICS			
GENDER (DSEX)	Civics student questionnaire	Which of the following best describes you?	Male (reference) Female
RACE/ETHNICITY (SRACE10)	Civics student questionnaire	Which of the following best describes you? Select one or more answer choices.	White (reference) Black Hispanic Asian American Indian/Alaskan Native Pacific Islander/Native Hawaiian 2+ Races and Other
LIMITED ENGLISH LANGUAGE PROFICIENCY (LEP)	School records	Student classified as English learner	No (reference) Yes
SOCIOECONOMIC STATUS INDICATORS			
PARENTS EDUCATION (PARED)	Civics student questionnaire	Parental education level (from two questions): (1) How far in school did your mother go? (2) How far in school did your father go?	Did not finish high school (reference) Graduated from high school Some education after high school Graduated from college
BOOKS IN THE HOME (B013801)	Civics student questionnaire	About how many books are there in your home?	Few (0-10) (reference) Enough to fill one shelf (11-25) Enough to fill one bookcase (26-100) Enough to fill several bookcases (more than 100)
SCHOOL NSLP STATUS (C051601)	Social studies school questionnaire	During this school year, about what percentage of students in your school was eligible to receive a free or reduced-price lunch through the National School Lunch Program?	1-25% (reference) 25-50% 51-75% 76-100% and all students
STUDENT NSLP STATUS (SLUNCH3)	School records		Not eligible (reference) Eligible

NOTE: NSLP = National School Lunch Program.
SOURCE: National Center for Education Statistics, NAEP Report Card: Civics (Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Dept. of Education, 2018), https://www.nationsreportcard.gov/civics/results/groups/.

Appendix Table C2: Civics Independent Variables List (Continued)

VARIABLE (NAEP ID)	QUESTIONNAIRE	ITEM OR INDEX	REGRESSION CATEGORIES
OPPORTUNITY TO LEARN VARIABLES			
TEACHER BACKGROUND/PREPARATION			
TEACHING EXPERIENCE (T127501)	Social studies teacher questionnaire	Excluding student teaching, how many years have you taught civics, geography, history, or social studies in grades 6 through 12, counting this year?	Less than 1 year (reference) 1-2 years 3-5 years 6-10 years 11-20 years 21 or more years
PROFESSIONAL DEVELOPMENT (T119901)	Social studies teacher questionnaire	During the last two years, did you participate in or lead any professional development activities related to the teaching of civics, geography, history, or social studies?	No (reference) Yes
POLITICAL SCIENCE DEGREE (T094201)	Social studies teacher questionnaire	Did you have a major, minor, or special emphasis in political science as part of your undergraduate coursework?	No (reference) Yes, a major Yes, a minor or special emphasis
CIVICS TEACHING ROLE (T137704)	Social studies teacher questionnaire	Which best describes your role in teaching social studies to this class? "I have primary responsibility for teaching civics and/or United States government."	No (reference) Yes
TEACHER SATISFACTION (T132701)	Social studies teacher questionnaire	How much does each of the following statements apply to you as a teacher? "I am satisfied with being a teacher at this school."	Not at all like me (reference) A little bit like me Somewhat like me Quite a bit like me Exactly like me
TEACHING EMPHASIS			
STUDENTS WRITING OPINION (T138202)	Social studies teacher questionnaire	During this school year so far, how often have you done each of the following as part of social studies instruction with this class? "Had students write about their opinion on a problem or issue (e.g., in a letter, e-mail, or blog post)"	Never (reference) Once Two or three times Four or five times More than five times
STUDENTS TAKING PART IN ROLE-PLAY/MOCK TRIAL (T138204)	Social studies teacher questionnaire	During this school year so far, how often have you done each of the following as part of social studies instruction with this class? "Had students take part in role-playing, mock trials, or dramas about social studies topics"	Never (reference) Once Two or three times Four or five times More than five times

NOTE: NSLP = National School Lunch Program.
SOURCE: National Center for Education Statistics, NAEP Report Card: Civics (Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Dept. of Education, 2018), https://www.nationsreportcard.gov/civics/results/groups/.

Appendix Table C2: Civics Independent Variables List (Continued)

VARIABLE (NAEP ID)	QUESTIONNAIRE	ITEM OR INDEX	REGRESSION CATEGORIES
STUDENTS USING COMPUTER TO CREATE REPORTS (T138303)	Social studies teacher questionnaire	When students in this class work on social studies, to what extent do they use computers or other digital devices to do the following? "Create reports or projects about social studies using different forms of media (e.g., a slide presentation that combines texts and video clips)"	Never or hardly ever (reference) Once in a while Sometimes Often Always or almost always
EMPHASIS ON DISCUSSING POLITICAL PROCESS (T138807)	Social studies teacher questionnaire	In your social studies class this school year, how much have you emphasized teaching your students each of the following? "Discussing the political process and government with others"	Not at all (reference) Very little Some Quite a bit A lot
EMPHASIS ON TAKING POSITION ON ISSUES (T139607)	Social studies teacher questionnaire	In your social studies class this school year, how much have you emphasized teaching your students each of the following? "Taking and defending a position about a historical issue (e.g., how changes in transportation have affected the United States economy)"	Not at all (reference) Very little Some Quite a bit A lot
STUDENT INTRA/INTERPERSONAL FACTORS			
CIVICS INTEREST/ ENJOYMENT INDEX (SQCATC4)	Civics student questionnaire	Students' interest/enjoyment in civics index Based on responses to five questions: How much does each of the following statements describe you? 1. Civics and/or United States government are my favorite topics to study. 2. I enjoy doing schoolwork about civics and/or United States government. 3. I enjoy discussing civics and/or United States government topics with others. 4. I think that civics and/or United States government schoolwork helps me understand what is happening in the world around me. 5. I think that learning about civics and/or United States government topics will be important for my future.	Low (reference) Moderate High

NOTE: NSLP = National School Lunch Program.
SOURCE: National Center for Education Statistics, NAEP Report Card: Civics (Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Dept. of Education, 2018), https://www.nationsreportcard.gov/civics/results/groups/.

Appendix Table C2: Civics Independent Variables List (Continued)

VARIABLE (NAEP ID)	QUESTIONNAIRE	ITEM OR INDEX	REGRESSION CATEGORIES
CIVICS COMMUNITY ENGAGEMENT INDEX (SQCATC8)	Civics student questionnaire	Students' views on community engagement index Based on responses to five questions: How much does each of the following statements describe you? 1. I think I can make a difference in my community. 2. I think being actively involved in community issues is my responsibility. 3. I think being concerned about state and local issues is an important responsibility for everybody. 4. I have good ideas for programs and projects that would help solve problems in my community. 5. I expect to be involved in improving my community three years from now.	Low (reference) Moderate High
IMPORTANCE OF ATTENTION TO GOVERNMENT (P813805)	Civics student questionnaire	Do you think that you would be able to do each of the following? "Explain why it is important to pay attention to the political process and government."	I definitely can't (reference) I probably can't Maybe I probably can I definitely can
IMPORTANCE OF PARTICIPATION IN GOVERNMENT (P813806)	Civics student questionnaire	Do you think that you would be able to do each of the following? "Explain why it is important for individuals to participate in the political process and government."	I definitely can't (reference) I probably can't Maybe I probably can I definitely can

NOTE: NSLP = National School Lunch Program.
SOURCE: National Center for Education Statistics, NAEP Report Card: Civics (Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Dept. of Education, 2018), https://www.nationsreportcard.gov/civics/results/groups/.

Appendix Table C3: U.S. History Independent Variables

VARIABLE (NAEP ID)	QUESTIONNAIRE	ITEM OR INDEX	REGRESSION CATEGORIES
DEMOGRAPHICS			
GENDER (DSEX)	U.S. history student questionnaire	Which of the following best describes you?	Male (reference) Female
RACE/ETHNICITY (SRACE10)	U.S. history student questionnaire	Which of the following best describes you? Select one or more answer choices.	White (reference) Black Hispanic Asian American Indian/Alaskan Native Pacific Islander/Native Hawaiian 2+ Races and Other
LIMITED ENGLISH LANGUAGE PROFICIENCY (LEP)	School records	Student classified as English learner	No (reference) Yes
SOCIOECONOMIC STATUS INDICATORS			
PARENTS EDUCATION (PARED)	U.S. history student questionnaire	Parental education level (from two questions): (1) How far in school did your mother go? (2) How far in school did your father go?	Did not finish high school (reference) Graduated from high school Some education after high school Graduated from college
BOOKS IN THE HOME (B013801)	U.S. history student questionnaire	About how many books are there in your home?	Few (0-10) (reference) Enough to fill one shelf (11-25) Enough to fill one bookcase (26-100) Enough to fill several bookcases (more than 100)
SCHOOL NSLP STATUS (C051601)	Social studies school questionnaire	During this school year, about what percentage of students in your school was eligible to receive a free or reduced-price lunch through the National School Lunch Program?	1-25% (reference) 25-50% 51-75% 76-100% and all students
STUDENT NSLP STATUS (SLUNCH3)	School records		Not eligible (reference) Eligible

NOTE: NSLP = National School Lunch Program.
SOURCE: National Center for Education Statistics, NAEP Report Card: U.S. History (Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Dept. of Education, 2018), https://www.nationsreportcard.gov/ushistory/results/groups/.

Appendix Table C3: U.S. History Independent Variables (Continued)

VARIABLE (NAEP ID)	QUESTIONNAIRE	ITEM OR INDEX	REGRESSION CATEGORIES
OPPORTUNITY TO LEARN VARIABLES			
TEACHER BACKGROUND/PREPARATION			
TEACHING EXPERIENCE (T127501)	Social studies teacher questionnaire	Excluding student teaching, how many years have you taught civics, geography, history, or social studies in grades 6 through 12, counting this year?	Less than 1 year (reference) 1-2 years 3-5 years 6-10 years 11-20 years 21 or more years
PROFESSIONAL DEVELOPMENT (T119901)	Social studies teacher questionnaire	During the last two years, did you participate in or lead any professional development activities related to the teaching of civics, geography, history, or social studies?	No (reference) Yes
HISTORY DEGREE (T077301)	Social studies teacher questionnaire	Did you have a major, minor, or special emphasis in history or history education as part of your undergraduate coursework?	No (reference) Yes, a major Yes, a minor or special emphasis
HISTORY TEACHING ROLE (T137706)	Social studies teacher questionnaire	Which best describes your role in teaching social studies to this class? "I have primary responsibility for teaching United States history."	No (reference) Yes
TEACHER SATISFACTION (T132701)	Social studies teacher questionnaire	How much does each of the following statements apply to you as a teacher? "I am satisfied with being a teacher at this school."	Not at all like me (reference) A little bit like me Somewhat like me Quite a bit like me Exactly like me
TEACHING EMPHASIS			
STUDENTS USE INTERNET FOR EVIDENCE (T138301)	Social studies teacher questionnaire	When students in this class work on social studies, to what extent do they use computers or other digital devices to do the following? "Use the Internet to look for evidence or sources (e.g., text documents, photographic images, or films)"	Never or hardly ever (reference) Once in a while Sometimes Often Always or almost always
ASSIGNMENT WITH MULTIPLE SOURCES (H813105)	U.S. history student questionnaire	In your social studies class this year, how often do you get the following assignments? "Responses to questions based on information from several sources (for example, letters, cartoons, or maps)"	Never or hardly ever (reference) Less than half of the lessons About half of the lessons More than half of the lessons All or almost all of the lessons

NOTE: NSLP = National School Lunch Program.
SOURCE: National Center for Education Statistics, NAEP Report Card: U.S. History (Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Dept. of Education, 2018), https://www.nationsreportcard.gov/ushistory/results/groups/.

Appendix Table C3: U.S. History Independent Variables (Continued)

VARIABLE (NAEP ID)	QUESTIONNAIRE	ITEM OR INDEX	REGRESSION CATEGORIES
ASSESS LONG WRITTEN RESPONSES (T138403)	Social studies teacher questionnaire	In your social studies class this year, how often do you use each of the following to assess student progress in social studies? "Long written response (e.g., several paragraphs)"	Never (reference) About once or twice a year About once or twice a month About once for twice a week Every day or almost every day
STUDENT INTRA/INTERPERSONAL FACTORS			
U.S. HISTORY INTEREST/ ENJOYMENT INDEX (SQCATH4)	U.S. history student questionnaire	U.S. history interest/enjoyment index Based on responses to five questions: How much does each of the following statements describe you? 1. United States history is one of my favorite subjects to study. 2. I enjoy doing schoolwork about United States history. 3. I enjoy discussing United States history topics with others. 4. I think that United States history schoolwork helps me understand what is happening in the world around me. 5. I think that learning about United States history topics will be important for my future.	Low (reference) Moderate High

NOTE: NSLP = National School Lunch Program.
SOURCE: National Center for Education Statistics, NAEP Report Card: U.S. History (Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Dept. of Education, 2018), https://www.nationsreportcard.gov/ushistory/results/groups/.

Appendix Table C3: U.S. History Independent Variables (Continued)

VARIABLE (NAEP ID)	QUESTIONNAIRE	ITEM OR INDEX	REGRESSION CATEGORIES
STUDENTS' CONFIDENCE IN U.S. HISTORY KNOWLEDGE/ SKILLS INDEX (SQCATH5)	U.S. history student questionnaire	Students' confidence in U.S. history knowledge/skills index Based on responses to six questions: How much does each of the following statements describe you? 1. Explain the causes and effects of important events in United States history Explain how time periods in United States history are similar or different 2. Judge whether information from a source (for example, text, visual graphics, video, or audio) is accurate 3. Understand the relationship between two historical events 4. Compare and evaluate different points of view about the past (for example, different ideas about what caused the Civil War) 5. Come up with research questions about why or how something happened in the past (for example, what were the causes of the Great Depression) 6. Take and defend a position about a historical issue (for example, how changes in transportation have affected the United States economy)	Low (reference) Moderate High
STUDENTS' PERSPECTIVE TAKING IN HISTORICAL THINKING INDEX (SQCATH8)	U.S. history student questionnaire	Students' perspective taking in historical thinking index Based on responses to five questions: How much does each of the following statements describe you? 1. I form opinions about historical events only after I have information from more than one source. 2. I need to know the history leading up to an event to truly understand it. 3. I want to know what lies behind the story when I study a conflict in history. 4. I try to understand others better by imagining how things look from their perspective. 5. I think that there is more than one side to every question, and I try to look at all of them.	Low (reference) Moderate High

NOTE: NSLP = National School Lunch Program.
SOURCE: National Center for Education Statistics, NAEP Report Card: U.S. History (Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Dept. of Education, 2018), https://www.nationsreportcard.gov/ushistory/results/groups/.

Appendix Table C4: Civics Regression Results

VARIABLE	MODEL 1			MODEL 2			MODEL 3			MODEL 4			MODEL 5		
	B	SE	β	B	SE	β	B	SE	β	B	SE	β	B	SE	β
Gender															
Female versus Male	1.732	1.095	1.897	1.773	1.079	1.913	1.711	1.116	1.910	1.132	1.371	1.552	1.454	1.332	1.936
Race/Ethnicity															
Black versus White	-14.358	1.781	-25.573**	-9.480	1.675	-15.876**	-9.388	1.645	-15.440**	-9.288	1.747	-16.226**	-8.619	1.882	-16.224**
Hispanic versus White	-7.853	1.505	-11.821**	-0.127	1.413	-0.179	-0.175	1.454	-0.254	-0.648	1.801	-1.167	-0.668	1.665	-1.112
Asian versus White	6.289	2.056	12.931**	8.607	1.934	16.643**	8.304	1.921	15.951**	8.233	2.120	17.458**	8.019	2.008	16.098**
American Indian/Alaska Native versus White	-1.955	8.080	-15.795	-1.363	7.164	-9.763	-1.316	7.332	-9.651	-1.185	5.765	-6.829	-1.548	4.331	-6.705
Pacific Islander/Native Hawaiian versus White	-2.738	5.454	-14.932**	-0.259	6.799	-1.759	-0.383	7.189	-2.756	-0.457	8.866	-4.053	-0.954	9.212	-8.786
2+ Races and Other versus White	-2.243	2.261	-5.070*	-0.664	2.626	-1.743	-0.626	2.754	-1.724	-0.750	3.338	-2.504	-0.906	3.415	-3.095
English Language Proficiency															
English Language Learner versus Native Speaker	-19.153	1.886	-36.122**	-12.056	2.189	-26.394**	-11.885	2.221	-26.402**	-10.293	2.635	-27.118**	-10.234	2.577	-26.378**
Parents Education															
Graduate H.S. versus Did Not Graduate H.S.	—	—	—	0.583	1.958	1.141	0.616	1.932	1.190	0.425	2.380	1.012	0.010	2.358	0.023
Post H.S. versus Did Not Graduate H.S.	—	—	—	6.635	1.977	13.114**	6.500	2.060	13.390**	5.339	2.408	12.856**	4.829	2.350	11.348**
Graduate College versus Did Not Graduate H.S.	—	—	—	7.331	1.637	12.003**	7.010	1.729	12.122**	5.733	2.110	12.095**	4.732	2.070	9.795**
Books in the home															
11-25 versus 0-10	—	—	—	4.083	1.675	6.840**	4.272	1.663	7.105**	3.407	2.034	6.932**	3.303	1.965	6.489**
26-100 versus 0-10	—	—	—	8.621	1.676	14.448**	8.188	1.755	14.372**	6.693	1.999	13.381**	6.098	2.000	12.198**
>100 versus 0-10	—	—	—	11.980	1.804	21.617**	11.216	1.896	21.270**	9.454	2.291	21.661**	8.420	2.287	19.257**
School NSLP Status															
26-50% versus 1-25%	—	—	—	-0.614	2.077	-1.275	-0.436	2.235	-0.974	-0.572	2.603	-1.490	-0.913	2.407	-2.197
51-75% versus 1-25%	—	—	—	-2.424	1.804	-4.371*	-2.379	1.836	-4.368**	-2.675	1.864	-4.987**	-2.939	1.752	-5.150**
76-100% and all students versus 1-25%	—	—	—	-1.992	1.672	-3.331*	-1.591	1.737	-2.763	-1.501	2.027	-3.043	-1.972	1.973	-3.889*
Student NSLP Status															
Eligible versus Not eligible	—	—	—	-7.412	1.197	-8.869**	-7.031	1.220	-8.576**	-5.823	1.617	-9.413**	-5.845	1.582	-9.246
Teaching Experience															
1-2 years versus Less than 1 year	—	—	—	—	—	—	0.875	3.030	2.653	1.232	3.492	4.301	1.717	3.270	5.616
3-5 years versus Less than 1 year	—	—	—	—	—	—	1.840	2.772	5.100	1.641	3.105	5.097	2.251	2.805	6.313*
6-10 years versus Less than 1 year	—	—	—	—	—	—	2.096	2.702	5.662**	2.075	3.047	6.323*	2.593	2.859	7.416*
11-20 years versus Less than 1 year	—	—	—	—	—	—	1.717	2.866	4.921	1.587	2.980	4.728	1.976	2.777	5.488*
21 or more versus Less than 1 year	—	—	—	—	—	—	1.555	3.158	4.900	1.992	3.187	6.348*	2.366	3.033	7.174*
Political Science Degree															
Major versus None	—	—	—	—	—	—	1.629	1.440	2.346	0.274	1.869	0.513	0.353	1.741	0.615
Minor versus None	—	—	—	—	—	—	-0.533	2.106	-1.122	-0.230	1.982	-0.456	0.043	2.021	0.086
Professional development															
Yes versus No	—	—	—	—	—	—	0.208	1.637	0.340	0.567	1.591	0.901	0.758	1.584	1.201
Teacher Satisfaction															
Little versus Not at all	—	—	—	—	—	—	0.728	4.606	3.354	0.682	4.666	3.181	0.899	4.323	3.888
Somewhat versus Not at all	—	—	—	—	—	—	0.993	3.978	3.949	1.058	4.598	4.866	1.102	4.287	4.726
Quite a Bit versus Not at all	—	—	—	—	—	—	0.621	3.981	2.471	0.572	4.592	2.626	—	0.659	4.205
Exactly versus Not at all	—	—	—	—	—	—	1.243	3.960	4.925	1.184	4.551	1.344	4.190	5.629	5.629
Civics Teaching Role															
Yes versus No	—	—	—	—	—	—	1.269	1.727	2.193	-0.027	1.731	-0.047	-0.049	1.680	-0.082

Statistical significance: * *p* < .05, ** *p* < .01.
NOTE: H.S. = high school, NSLP = National School Lunch Program.

Appendix Table C4: Civics Regression Results (Continued)

VARIABLE	MODEL 1 B	MODEL 1 SE	MODEL 1 β	MODEL 2 B	MODEL 2 SE	MODEL 2 β	MODEL 3 B	MODEL 3 SE	MODEL 3 β	MODEL 4 B	MODEL 4 SE	MODEL 4 β	MODEL 5 B	MODEL 5 SE	MODEL 5 β
Students writing opinion															
Once versus Never	—	—	—	—	—	—	—	—	—	-2.278	3.651	-8.320*	-2.689	3.829	-10.298**
Two or three times versus Never	—	—	—	—	—	—	—	—	—	-0.963	3.317	-3.196	-1.533	3.368	-5.163
Four or five times versus Never	—	—	—	—	—	—	—	—	—	-1.519	3.271	-4.968	-1.835	3.400	-6.238
More than five times versus Never	—	—	—	—	—	—	—	—	—	-0.568	3.385	-1.924	-1.047	3.483	-3.647
Students taking part in role-play/mock-trial															
Once versus Never	—	—	—	—	—	—	—	—	—	1.373	2.183	2.997	1.407	2.018	2.838
Two or three times versus Never	—	—	—	—	—	—	—	—	—	0.716	1.708	1.224	0.701	1.691	1.186
Four or five times versus Never	—	—	—	—	—	—	—	—	—	3.099	2.357	7.304**	2.823	2.345	6.619**
More than five times versus Never	—	—	—	—	—	—	—	—	—	0.144	3.549	0.512	—	0.215	3.308
Students using computer to create reports															
Once in while versus Never	—	—	—	—	—	—	—	—	—	2.106	2.416	5.087*	2.259	2.219	5.014*
Sometimes versus Never	—	—	—	—	—	—	—	—	—	1.691	2.430	4.107	1.737	2.374	4.123
Often versus Never	—	—	—	—	—	—	—	—	—	1.251	2.389	2.989	1.320	2.307	3.046
Always or almost always versus Never	—	—	—	—	—	—	—	—	—	2.344	3.048	7.145*	2.438	2.997	7.309*
Emphasis on discussing political process															
Very little versus Not at all	—	—	—	—	—	—	—	—	—	-0.317	2.919	-0.924	-0.770	2.751	-2.119
Some versus Not at all	—	—	—	—	—	—	—	—	—	1.283	2.429	3.116	0.960	2.339	2.244
Quite a bit versus Not at all	—	—	—	—	—	—	—	—	—	2.237	2.557	5.720*	1.974	2.446	4.828*
A lot versus Not at all	—	—	—	—	—	—	—	—	—	1.978	2.321	4.590*	1.619	2.307	3.734
Emphasis on taking position on issues															
Very little versus Not at all	—	—	—	—	—	—	—	—	—	-0.025	2.784	-0.071	0.455	2.693	1.226
Some versus Not at all	—	—	—	—	—	—	—	—	—	-0.762	2.001	-1.524	-0.496	2.087	-1.036
Quite a bit versus Not at all	—	—	—	—	—	—	—	—	—	-0.253	2.437	-0.617	-0.027	2.441	-0.065
A lot versus Not at all	—	—	—	—	—	—	—	—	—	0.884	2.622	2.318	0.722	2.688	1.940
Civics Interest/Enjoyment Index															
Moderate versus low	—	—	—	—	—	—	—	—	—	—	—	—	-1.790	1.670	-2.990
High versus low	—	—	—	—	—	—	—	—	—	—	—	—	-1.682	2.066	-3.477
Moderate versus low	—	—	—	—	—	—	—	—	—	—	—	—	-2.151	1.306	-2.810*
High versus low	—	—	—	—	—	—	—	—	—	—	—	—	-3.914	1.915	-7.495**
Importance of attention to government															
I probably can't versus I definitely can't	—	—	—	—	—	—	—	—	—	—	—	—	-1.509	3.139	-4.736
Maybe versus I definitely can't	—	—	—	—	—	—	—	—	—	—	—	—	-1.185	3.129	-3.708
I probably can versus I definitely can't	—	—	—	—	—	—	—	—	—	—	—	—	-0.007	3.198	-0.024
I definitely can versus I definitely can't	—	—	—	—	—	—	—	—	—	—	—	—	1.006	3.916	3.939
Importance of participation in government															
I probably can't versus I definitely can't	—	—	—	—	—	—	—	—	—	—	—	—	2.347	3.077	7.220
Maybe versus I definitely can't	—	—	—	—	—	—	—	—	—	—	—	—	2.490	2.910	7.246
I probably can versus I definitely can't	—	—	—	—	—	—	—	—	—	—	—	—	4.788	3.171	15.182
I definitely can versus I definitely can't	—	—	—	—	—	—	—	—	—	—	—	—	4.808	3.787	18.208
R^2		0.177			0.303			0.308			0.350			0.387	

Statistical significance: * $p < .05$, ** $p < .01$.
NOTE: H.S. = high school, NSLP = National School Lunch Program.

Appendix Table C5: U.S. History Regression Results

VARIABLE	MODEL 1			MODEL 2			MODEL 3			MODEL 4			MODEL 5		
	B	SE	β	B	SE	β	B	SE	β	B	SE	β	B	SE	β
Gender															
Female versus Male	-4.216	0.871	-3.673**	-4.332	0.913	-3.954**	-4.582	0.899	-4.120**	-4.784	0.888	-4.248**	-4.989	0.892	-4.452**
Race/Ethnicity															
Black versus White	-16.936	1.432	-24.257**	-6.999	1.729	-12.102**	-6.666	1.753	-11.683**	-6.945	1.722	-11.959**	-7.198	1.713	-12.333**
Hispanic versus White	-11.420	1.131	-12.917**	0.942	1.166	1.098	0.849	1.156	0.981	0.431	1.182	0.510	0.302	1.104	0.333
Asian versus White	3.965	1.880	7.455**	5.288	1.806	9.552**	5.033	1.857	9.345**	4.865	1.842	8.962	4.770	1.774	8.461**
American Indian/Alsk Native versus White	-3.774	3.100	-11.697**	-0.765	2.852	-2.181	-0.535	2.768	-1.481	-0.701	2.845	-1.995	-0.504	2.786	-1.404
Pacific Islander/Native Hawaiian versus White	-1.589	6.612	-10.506	-1.076	6.442	-6.933	-1.254	6.169	-7.734	-1.349	6.404	-8.640	-1.044	6.117	-6.388
2+ Races and Other versus White	-2.794	2.233	-6.237**	-0.644	2.555	-1.646	-0.589	2.585	-1.522	-0.716	2.542	-1.820	-0.730	2.481	-1.810
English Language Proficiency															
English Language Learner versus Native Speaker	-22.015	1.564	-34.439	-15.454	1.710	-26.430**	-14.881	1.737	-25.846**	-14.922	1.708	-25.492**	-14.663	1.687	-24.734**
Parents Education															
Graduate H.S. versus Did Not Graduate H.S.	—	—	—	1.902	1.480	2.816	1.704	1.529	2.605	1.660	1.508	2.504	1.282	1.554	1.992
Post H.S. versus Did Not Graduate H.S.	—	—	—	6.705	1.514	10.150**	6.320	1.627	10.286**	6.057	1.630	9.873**	5.482	1.569	8.601**
Graduate College versus Did Not Graduate H.S.	—	—	—	7.882	1.405	11.076**	7.539	1.431	10.789**	7.247	1.434	10.394**	6.642	1.411	9.369**
Books in the home															
11-25 versus 0-10	—	—	—	3.048	1.313	4.003**	3.187	1.354	4.315**	3.003	1.364	4.096**	2.762	1.355	3.741**
26-100 versus 0-10	—	—	—	9.838	1.167	11.485**	9.651	1.177	11.356**	9.319	1.185	11.045**	7.936	1.190	9.445**
>100 versus 0-10	—	—	—	12.333	1.557	19.205**	11.909	1.614	19.225**	11.819	1.612	19.057**	10.006	1.636	16.372**
School NSLP Status															
26-50% versus 1-25%	—	—	—	1.473	1.942	2.861	1.530	1.840	2.815	1.460	1.927	2.813	1.352	2.007	2.713
51-75% versus 1-25%	—	—	—	-0.829	1.353	-1.122	-1.044	1.298	-1.356	-1.029	1.348	-1.387	-1.089	1.303	-1.419
76-100% and all students versus 1-25%	—	—	—	-2.179	1.728	-3.764*	-1.942	1.652	-3.208	-1.942	1.599	-3.106	-2.102	1.612	-3.388*
Student NSLP Status															
Eligible versus Not eligible	—	—	—	-7.921	1.188	-9.413**	-7.544	1.205	-9.094**	-7.294	1.218	-8.882**	-7.347	1.154	-8.481**
Teaching Experience															
1-2 years versus Less than 1 year	—	—	—	—	—	—	0.917	2.192	2.010	0.927	2.182	2.022	0.944	2.291	2.162
3-5 years versus Less than 1 year	—	—	—	—	—	—	1.259	2.123	2.674	1.262	2.140	2.701	1.612	2.231	3.596
6-10 years versus Less than 1 year	—	—	—	—	—	—	2.880	2.150	6.191**	2.821	2.162	6.100**	3.010	2.211	6.654**
11-20 years versus Less than 1 year	—	—	—	—	—	—	2.488	2.245	5.587*	2.474	2.268	5.610*	2.698	2.334	6.297**
21 or more versus Less than 1 year	—	—	—	—	—	—	2.304	2.083	4.799*	2.372	2.086	4.947*	2.537	2.162	5.484**
Teacher Satisfaction															
Little versus Not at all	—	—	—	—	—	—	1.036	4.087	4.234	1.023	4.001	4.091	0.763	3.842	2.931
Somewhat versus Not at all	—	—	—	—	—	—	1.254	1.254	3.261	4.091	1.227	3.164	3.883	1.237	3.079
Quite a Bit versus Not at all	—	—	—	—	—	—	0.862	0.862	0.862	3.363	2.898	0.927	3.213	2.977	0.976
Exactly versus Not at all	—	—	—	—	—	—	1.384	3.135	4.339	1.465	2.981	4.366	1.406	2.908	4.087
History Degree															
Major versus None	—	—	—	—	—	—	-0.829	1.045	-0.866	-0.755	1.049	-0.792	-0.482	1.048	-0.505
Minor versus None	—	—	—	—	—	—	0.171	1.468	0.251	0.169	1.440	0.243	0.164	1.381	0.226
History Teaching Role															
Yes versus No	—	—	—	—	—	—	2.273	1.145	2.602*	2.193	1.131	2.479*	2.122	1.104	2.343*
Professional development															
Yes versus No	—	—	—	—	—	—	2.811	1.097	3.085	2.682	1.111	2.981	2.652	1.059	2.808

Statistical significance: * $p < .05$, ** $p < .01$.
NOTE: H.S. = high school, NSLP = National School Lunch Program.

Appendix Table C5: U.S. History Regression Results (Continued)

VARIABLE	MODEL 1			MODEL 2			MODEL 3			MODEL 4			MODEL 5		
	B	SE	β	B	SE	β	B	SE	β	B	SE	β	B	SE	β
Students use internet for evidence															
Once in while versus Never	—	—	—	—	—	—	—	—	—	0.109	2.055	0.225	0.267	1.936	0.517
Sometimes versus Never	—	—	—	—	—	—	—	—	—	-0.497	2.125	-1.057	-0.413	2.032	-0.840
Often versus Never	—	—	—	—	—	—	—	—	—	-0.676	2.212	-1.494	-0.555	2.110	-1.172
Always or almost always versus Never	—	—	—	—	—	—	—	—	—	0.050	2.357	0.118	0.018	2.226	0.039
Assignment with multiple sources															
<Half of the lessons versus no lessons	—	—	—	—	—	—	—	—	—	4.274	1.233	5.269**	2.839	1.135	3.221**
About half lessons versus no lessons	—	—	—	—	—	—	—	—	—	3.666	1.282	4.698**	1.431	1.209	1.730
>Half of the lessons versus no lessons	—	—	—	—	—	—	—	—	—	7.435	1.080	8.027**	3.324	1.105	3.673**
All/most all lessons versus no lessons	—	—	—	—	—	—	—	—	—	3.937	1.268	4.993**	-0.681	1.322	-0.901
Assess long written responses															
Once/twice a year versus never	—	—	—	—	—	—	—	—	—	0.666	2.157	1.435	0.501	2.144	1.075
Once/twice a month versus never	—	—	—	—	—	—	—	—	—	0.577	1.971	1.136	0.520	1.988	1.033
Once/twice a week versus never	—	—	—	—	—	—	—	—	—	0.393	2.448	0.963	0.268	2.454	0.657
Every day or almost versus never	—	—	—	—	—	—	—	—	—	-0.014	2.848	-0.040	0.032	3.055	0.098
U.S. History Interest/Enjoyment Index															
Moderate versus low	—	—	—	—	—	—	—	—	—	—	—	—	-3.493	1.029	-3.594**
High versus low	—	—	—	—	—	—	—	—	—	—	—	—	-4.017	1.191	-4.786**
Students' confidence in U.S. history knowledge/skills Index															
Moderate versus low	—	—	—	—	—	—	—	—	—	—	—	—	6.177	1.107	6.838**
High versus low	—	—	—	—	—	—	—	—	—	—	—	—	12.850	1.053	13.531**
Students' perspective taking in historical thinking index															
Moderate versus low	—	—	—	—	—	—	—	—	—	—	—	—	1.809	1.034	1.870
High versus low	—	—	—	—	—	—	—	—	—	—	—	—	5.322	1.203	6.402**
R²		0.193			0.314			0.322			0.330			0.362	

Statistical significance: * p < .05, ** p < .01.
NOTE: H.S. = high school, NSLP = National School Lunch Program.

About the Authors

Madeline J. Goodman is a researcher and author for the National Assessment of Educational Progress (NAEP) at ETS. Since joining the organization, Dr. Goodman has served as coordinator for the NAEP social science assessments and director for the NAEP website. Dr. Goodman has written and co-authored numerous national, state, district, and special reports based on national and international assessments including, most recently, *Opportunity Across the States* (2021), *Too Big to Fail: Millennials on the Margins* (2018), and *America's Skills Challenge: Millennials and the Future* (2015). Dr. Goodman received a Ph.D. in U.S. history from Carnegie Mellon University and is the recipient of a Spencer Fellowship for research on improvement in education and a Fulbright Scholar Award.

Andrew R. Weiss is an assessment development manager at ETS. He has contributed to numerous testing programs, including the *Advanced Placement®*, *PRAXIS®*, and *CLEP®* programs, and for much of the last 20 years he has led test development for NAEP social sciences assessments in U.S. history, civics, geography, and economics. Weiss was coauthor of an early study of applying technology to assessment, *Problem Solving in Technology-Rich Environments* (2007), as well as several reports analyzing the results of NAEP science and social sciences assessments. Recently he has appeared on several panels exploring how assessments can be used to improve civic learning. Weiss is ABD in U.S. history from Cornell University and served as a lecturer for several years at Princeton University.

Jesse R. Sparks is a senior research scientist in the Learning and Assessment Foundations and Innovations Center in the Research and Development Division at ETS. Since joining the organization, Dr. Sparks has led the development of multiple cognitively based assessment frameworks and interactive task designs for measuring research and inquiry skills and practices in K–12 English language arts, science, and social science domains. This includes leading a NAEP Survey Assessment Innovations Lab (SAIL) project to develop an interactive virtual world for assessing middle school students' multiple-document inquiry processes and skills and developing conceptual frameworks and learning progression-based assessments of research, inquiry, and argumentation skills under the Cognitively Based Assessment *of, for,* and *as* Learning (*CBAL*®) research initiative at ETS. Dr. Sparks has published numerous refereed articles, book chapters, and research reports. Dr. Sparks received a PhD in learning sciences and a certificate in cognitive science from Northwestern University.

Kelsey D. Dreier is an education data analyst who specializes in the management, analysis, reporting and visualization of key metrics to support research and data-driven programs. Her research interests include education policy, psychometrics, and literacy across the lifetime. She currently serves as the data specialist for The Campagna Center, a not-for-profit organization in Alexandria, VA, that provides educational and social development programs for children, teens, and adults. Before joining The Campagna Center, Dreier served as a senior research assistant in ETS's Learning & Assessment Foundations group within the Research and Development Division. At ETS, she worked within the Center for Global Assessment on the Programme for International Student Assessment (PISA) and on policy reports for the ETS Center for Research on Human Capital and Education. She has also assisted with several Institute for Education Sciences grants. Dreier earned an Ed.M. in educational statistics, measurement, and evaluation at Rutgers University and a Bachelor of Science degree in cognitive studies and psychology with a minor in quantitative studies from Vanderbilt University. Before joining ETS, Dreier worked with a wide range of research divisions at Vanderbilt including the Institute of Public Health: Qualitative Research Core, the Education and Brain Research Lab, and the Kennedy Center Reading Clinic

Endnotes

1 Mitja Sardoc, "Democratic Education at 30: An interview with Dr. Amy Gutmann," *Theory and Research in Education* 16, no. 2 (2018): 244–52, https://doi.org/10.1177/1477878518774087.

2 See, for example, John Dewey, *Democracy and Education* (New York: Macmillan, 2016); Carl F. Kaestle, *Pillars of the Republic: Common Schools and American Society, 1780–1860* (New York: Hill and Wang, 1983); Ira Katznelson and Margaret Weir, *Schooling for All: Class, Race, and the Decline of the Democratic Ideal* (Berkeley, CA: University of California Press, 1988); David B. Tyack, *The One Best System: A History of American Urban Education* (Cambridge, MA: Harvard University Press, 1974); Donald Warren, "Original Intents: Public Schools as Civic Education," *Theory into Practice* 27, no. 4 (1988): 243–49, https://doi.org/10.1080/00405848809543360.

3 David P. Gardner, Yvonne W. Larsen, William O. Baker, Ann Campbell, Emeral A. Crosby, Charles A. Foster, Jr., Norman C. Francis, et al. *A Nation at Risk: The Imperative for Educational Reform. An Open Letter to the American People* (Washington, DC: Government Printing Office, 1983), 7.

4 Education for Democracy Project, *Education for Democracy: A Statement of Principles, Guidelines for Strengthening the Teaching of Democratic Values* (Washington, DC: American Federation of Teachers, 1987); Center for Civic Education, *Civitas: A Framework for Civic Education* (Calabasas, CA: Center for Civic Education, 1991); Center for Civic Education, *National Standards for Civics and Government* (Calabasas, CA: Center for Civic Education, 1994); National Assessment Governing Board, *Civics Framework for the 1998 National Assessment of Educational Progress: NAEP Civics Consensus Project* (Washington, DC: National Assessment Governing Board, 1996); Judith Torney-Purta and Susan Vermeer Lopez, *Developing Citizenship Competencies from Kindergarten through Grade 12: A Background Paper for Policymakers and Educators* (Denver, CO: Education Commission of the States, 2006).

5 Educating for American Democracy, *The Roadmap to Educating for American Democracy* (Cambridge, MA: iCivics, 2021), https://www.educatingforamericandemocracy.org/the-roadmap/; Carol D. Lee, Gregory White, and Dian Dong, eds., *Educating for Civic Reasoning and Discourse* (Washington DC: National Academy of Education, 2021), https://naeducation.org/educating-for-civic-reasoning-and-discourse/; National Council for the Social Studies, *The College, Career, & Civic Life (C3) Framework for Social Studies State Standards: Guidance for Enhancing the Rigor of K-12 Civics, Economics, Geography, and History* (Silver Spring, MD: National Council for the Social Studies, 2013), https://www.socialstudies.org/sites/default/files/c3/C3-Framework-for-Social-Studies.pdf; Laura Hamilton and Ace Parsi, *Monitoring Civic Learning Opportunities and Outcomes: Lessons from an ETS/EAD Symposium* (Princeton, NJ: ETS, 2022), https://www.ets.org/Media/Research/pdf/Research_Notes_Hamilton_Parsi.pdf.

6 Lee et al., *Educating for Civic Reasoning and Discourse*; National Council for the Social Studies, *C3 Framework.*

7 Jan Brennan, *ESSA: Mapping Opportunities for Civic Education* (Denver, CO: Education Commission of the States, 2017); Jonathan Gould, ed., *Guardian of Democracy: The Civic Mission of Schools* (Philadelphia, PA: University of Pennsylvania, Leonore Annenberg Institute for Civics, 2011); Lisa Guilfoile and Brady Delander, *Guidebook*: *Six Proven Practices for Effective Civic Learning* (Denver, CO: Education Commission of the States, 2014), http://www.ecs.org/clearinghouse/01/10/48/11048.pdf; Steven Hart and Christian Wandeler, "The Impact of Action Civics Service-Learning on Eighth-Grade Students' Civic Outcomes," *International Journal of Research on Service-Learning and Community Engagement* 6, no. 1 (2018): 1–17, https://doi.org/10.37333/001c.6878.

8 Lee et al., *Educating for Civic Reasoning and Discourse*, 8.

9 Educating for American Democracy, *Roadmap;* National Council for the Social Studies, *C3 Framework.*

10 Raj Vinnakota, *From Civic Education to a Civic Learning Ecosystem: A Landscape Analysis and Case for Collaboration* (Red & Blue Works, 2019), https://rbw.civic-learning.org/wp-content/uploads/2019/12/CE_online.pdf; Keith C. Barton, "Shared Principles in History and Social Science Education," in *Palgrave Handbook of Research in Historical Culture and Education,* eds. Mario Carretero, Stefan Berger, and Maria Grever (London: Palgrave Macmillan, 2017), 449–67, https://doi.org/10.1057/978-1-137-52908-4_24; Michael E. Morrell, "Empathy and Democratic Education," *Public Affairs Quarterly* 21, no. 4 (2007): 381–403, https://www.jstor.org/stable/40441496; Mary E. Hylton, "The Role of Civic Literacy and Social Empathy on Rates of Civic Engagement among University Students," *Journal of Higher Education Outreach and Engagement* 22, no. 1 (2018): 87–106; Keith C. Barton and Linda S. Levstik, *Teaching History for the Common Good* (New York: Routledge, 2004); Jason Endacott and Sarah Brooks, "An Updated Theoretical and Practical Model for Promoting Historical Empathy," *Social Studies Research and Practice* 8, no. 1 (2013): 41–58, https://doi.org/10.1108/SSRP-01-2013-B0003.

11 Educating for American Democracy, Roadmap, 9; Peter N. Stearns, "Why Study History?" American Historical Association, 1998, https://www.historians.org/about-aha-and-membership/aha-history-and-archives/historical-archives/why-study-history-(1998).

12   James W. Pellegrino, Louis V. DiBello, and Susan R. Goldman, "A Framework for Conceptualizing and Evaluating the Validity of Instructionally Relevant Assessments," *Educational Psychologist* 51, no. 1 (2016): 59–81, https://doi.org/10.1080/00461520.2016.1145550. See also Susan R. Goldman, M. Anne Britt, Willard Brown, Gayle Cribb, MariAnne George, Cynthia Greenleaf, Carol D. Lee, Cynthia Shanahan, and Project READi, "Disciplinary Literacies and Learning to Read for Understanding: A Conceptual Framework for Disciplinary Literacy," *Educational Psychologist* 51, no. 2 (2016): 219–46, https://doi.org/10.1080/00461520.2016.1168741.

13   [13] Derek W. Black, "The Constitutional Compromise to Guarantee Education," *Stanford Law Review* 70 (2018), https://review.law.stanford.edu/wp-content/uploads/sites/3/2018/03/70-Stan.-L.-Rev.-735.pdf; Gary B., Jessie K., Cristopher R., Isaias R., Esmeralda V., Paul M., and Jaime R. v. Whitmer, 120 LRP 13567 (2020), https://static1.squarespace.com/static/5d7d1489b706a509a5b55f97/t/5ea8eaa77c46e83da091986c/1588128426333/Sixth+Cir.pdf.

14   Black, "Constitutional Compromise"; Gary B. et al. v. Whitmer, 19.

15   Patricia A. Alexander, Daniel L. Dinsmore, Emily Fox, Emily M. Grossnickle, Sandra M. Loughlin, Liliana Maggioni, Meghan M. Parkinson, and Fielding I. Winters, "Higher Order Thinking and Knowledge: Domain-General and Domain-Specific Trends and Future Directions," in *Assessment of Higher Order Thinking Skills: Current Perspectives on Cognition, Learning and Instruction*, eds. Gregory Schraw and Daniel H. Robinson (Information Age Publishing, 2011)*,* 47–88; Kadriye Ercikan and Peter Seixas, "Assessment of Higher Order Thinking: The Case of Historical Thinking," in *Assessment of Higher Order Thinking Skills,* 245–61.

16   "NAEP Achievement Levels," National Assessment Governing Board, accessed January 23, 2023, https://www.nagb.gov/focus-areas/NAEP-achievement-levels.html.

17   M. Anne Britt, Jean-François Rouet, and Amanda M. Durik, *Literacy Beyond Text Comprehension: A Theory of Purposeful Reading* (New York: Routledge, 2018).

18   Timothy Shanahan and Cynthia Shanahan, "Teaching Disciplinary Literacy to Adolescents: Rethinking Content-Area Literacy," *Harvard Educational Review* 78, no. 1 (2008): 40–59, https://doi.org/10.17763/haer.78.1.v62444321p602101; Cynthia Shanahan, Timothy Shanahan, and Cynthia Misichia, "Analysis of Expert Readers in Three Disciplines: History, Mathematics, and Chemistry," *Journal of Literacy Research* 43, no. 4 (2011): 393–429, https://doi.org/10.1177/1086296X11424071.

19   Goldman et al., "Disciplinary Literacies."

20   Lee et al., *Educating for Civic Reasoning and Discourse,* 1.

21   For a detailed discussion of how the NAEP assessments are designed, see Stephen Lazer, "A Large-Scale Assessment of Historical Knowledge and Reasoning: NAEP U.S. History," in New Directions in Assessing Historical Thinking, eds. Kadriye Ercikan and Peter Seixas (New York: Routledge, 2015).

22   "About the NAEP U.S. History Assessment," The Nation's Report Card, accessed January 23, 2023, https://www.nationsreportcard.gov/ushistory/about/digital-assessment/

23   Lee et al., *Educating for Civic Reasoning and Discourse.*

24   Douglas S. Massey, "Inheritance of Poverty or Inheritance of Place? The Emerging Consensus on Neighborhoods and Stratification," *Contemporary Sociology* 42*,* no. 5 (2013): 609–95, https://doi.org/10.1177/0094306113499534f; Sean F. Reardon, Ericka S. Weathers, Erin M. Fahle, Heewon Jang, and Demetra Kalogrides, *Is Separate Still Unequal? New Evidence on School Segregation and Racial Academic Achievement Gaps,* CEPA Working Paper No. 19-06 (Stanford, CA: Stanford Center for Education Policy Analysis, 2019), https://cepa.stanford.edu/sites/default/files/wp19-06-v092019.pdf.

25   For a brief discussion of the role of some OTL factors in relation to demographic characteristics alone, see Kei Kawashima-Ginsberg, *Do Discussion, Debate, and Simulations Boost NAEP Civics Performance?* CIRCLE Fact Sheet (Medford, MA: Center for Information and Research on Civic Learning and Engagement, April 2013), https://www.miciviced.org/wp-content/uploads/2018/10/circle_naepbechtelfactsheetapril30.final_.pdf.

26   The NAEP civics scale at grade 8 ranges from 0 to 300 (although student performance typically ranges from 110 to 190, representing the 10th and 90th percentiles, respectively). In 1998, the mean of the civics scale was set at 150 and the standard deviation at 35. The NAEP U.S. history scale at grade 8 ranges from 0 to 500 (although student performance typically ranges from 220 to 300, representing the 10th and 90th percentiles, respectively). In 1994, the mean of the U.S. history scale was set at 250 and the standard deviation at 50.

27   National Assessment Governing Board, *What Are NAEP Achievement Levels and How Are They Determined?* https://www.nagb.gov/content/dam/nagb/en/documents/naep/achievement-level-one-pager-4.6.pdf.

28 National Assessment Governing Board, *NAEP Achievement Levels,* 2.

29 For examples of item maps in NAEP civics, see https://www.nationsreportcard.gov/
itemmaps/?subj=CIV&grade=8&year=2018, and in NAEP U.S. history, see https://www.nationsreportcard.gov/
itemmaps/?subj=HIS&grade=8&year=2018.

30 See https://www.nationsreportcard.gov/civics/results/groups/ for NAEP civics results by subgroup and
https://www.nationsreportcard.gov/ushistory/results/groups/ for NAEP U.S. history results by subgroup.

31 Massey, "Inheritance of Poverty or Inheritance of Place?"; Reardon et al., *Is Separate Still Unequal?*

32 This narrowing of the White–Hispanic gap was also observed in the previous assessment year, 2014, when
compared to 1998. The White–Hispanic score gap also narrowed when comparing grade 8 reading results from
2019 with those of 1998. For the score gap results for civics, see https://www.nationsreportcard.gov/civics/results/
groups/.

33 American Academy of Arts and Sciences, *Our Common Purpose: Reinventing American Democracy for the 21st Century*
(Cambridge, MA: American Academy of Arts and Sciences, 2020); Equity in Civic Education Project, *Equity in Civic
Education* (Cambridge, MA: iCivics, 2020), https://civxnow.org/wp-content/uploads/2021/08/Equity-in-Civic-Educa
tion-White-Paper_Public.pdf; Tina L. Heafner and Paul G. Fitchett, "An Opportunity to Learn U.S. History: What
NAEP Data Suggest Regarding the Opportunity Gap," *The High School Journal* 98, no. 3 (2015): 226–49; Kei
Kawashima-Ginsberg, "The Future of Civic Education," *State Education Standard* 16, no. 3 (2016): 13–9 https://nas
be.nyc3.digitaloceanspaces.com/2016/09/NASBE-Standard-Sep16-FINALCompressed.pdf; Joseph Kahne and Ellen
Middaugh, *Democracy for Some: The Civic Opportunity Gap in High School*, Circle Working Paper 59 (Medford, MA:
Center for Information and Research on Civic Learning and Engagement, 2008), https://circle.tufts.edu/sites/de
fault/files/2019-12/WP59_TheCivicOpportunityGapinHighSchool_2008.pdf.

34 Kahne and Middaugh, *Democracy for Some*.

35 Eric A. Hanushek, Paul E. Peterson, Laura M. Talpey, and Ludger Woessmann, *Long-Run Trends in the U.S. SES-
Achievement Gap*, Working Paper no. 26764 (Washington, DC: National Bureau of Economic Research, 2020),
https://doi.org/10.3386/w26764. See also Markus Broer, Yifan Bai, and Fran Fonseca, *Socioeconomic Inequality and
Educational Outcomes: Evidence from Twenty Years of TIMSS* (New York: Springer Nature, 2019); National Center for
Education Statistics, *Improving the Measurement of Socioeconomic Status for the National Assessment of Educational
Progress: A Theoretical Foundation* (Washington, DC: National Center for Education Statistics, 2012),
https://nces.ed.gov/nationsreportcard/pdf/researchcenter/socioeconomic_factors.pdf

36 Massey, "Inheritance of Poverty or Inheritance of Place?"; Reardon et al., *Is Separate Still Unequal?*

37 Stephen N. Elliott, "Measuring Opportunity to Learn and Achievement Growth: Key Research Issues With
Implications for the Effective Education of All Students," *Remedial and Special Education* 36, no. 1 (2015): 58–64,
https://doi.org/10.1177/0741932514551282; Andrew C. Porter and Adam Gamoran, eds., *Methodological Advances in
Cross-National Surveys of Educational Achievement* (Washington, DC: National Academies Press, 2002).

38 Heafner and Fitchett, "An Opportunity to Learn U.S. History"; Johannes König, Sarantis Tachtsoglou, Sandra
Lammerding, Sarah Strauß, Günter Nold, and Andreas Rohde, "The Role of Opportunities to Learn in Teacher
Preparation for EFL Teachers' Pedagogical Content Knowledge," *The Modern Language Journal* 101, no. 1 (2017):
109–27, https://doi.org/10.1111/modl.12383.

39 Julie Cohen and Rebekah Berlin, "What Constitutes an 'Opportunity to Learn' in Teacher Preparation?" *Journal of
Teacher Education* 71, no. 4 (2020), 434–48, https://doi.org/10.1177/0022487119879893; Porter and Gamoran,
Methodological Advances in Cross-National Surveys of Educational Achievement.

40 Lorin W. Anderson, "Curricular Alignment: A Re-Examination," *Theory Into Practice* 41, no. 4 (2002): 255–60,
https://doi.org/10.1207/s15430421tip4104_9; Lorraine M. McDonnell, "Opportunity to Learn as a Research Concept
and a Policy Instrument," *Educational Evaluation and Policy Analysis* 17, no. 3 (1995): 305–22, https://doi.org/10.3102/
01623737017003305.

41 Susan De La Paz, "Effects of Historical Reasoning Instruction and Writing Strategy Mastery in Culturally and
Academically Diverse Middle School Classrooms," *Journal of Educational Psychology* 97, no. 2 (2005), 139–56,
https://doi.org/10.1037/0022-0663.97.2.139; Heafner and Fitchett, "An Opportunity to Learn U.S. History";
Kawashima-Ginsberg, *Do Discussion, Debate, and Simulations Boost NAEP Civics Performance?*

42 Joshua Littenberg-Tobias, "Teaching Citizens: What Can NAEP Civics Tell Us about Active Learning in Civics?"
Peabody Journal of Education 96, no. 3 (2021): 247–60, https://doi.org/10.1080/0161956X.2021.1942704; Daniel P.
Mayer, "Measuring Instructional Practice: Can Policymakers Trust Survey Data?" *Educational Evaluation and Policy*

Analysis 21, no. 1 (1999): 29–45; Mareike Kunter and Jürgen Baumert, "Who Is the Expert? Construct and Criteria Validity of Student and Teacher Ratings of Instruction," *Learning Environments Research* 9, no. 3 (2006): 231–51, https://doi.org/10.1007/s10984-006-9015-7.

43 Mathilde Almlund, Angela L. Duckworth, James Heckman, and Tim Kautz, "Personality Psychology and Economics," in *Handbook of the Economics of Education*, Volume 4, eds. Eric A. Hanushek, Stephen Machin, and Ludger Woessmann (Amsterdam: Elsevier, 2011): 4–182; Camille A. Farrington, Melissa Roderick, Elaine Allensworth, Jenny Nagaoka, Tasha Seneca Keyes, David W. Johnson, and Nicole O. Beechum, *Teaching Adolescents to Become Learners: The Role of Noncognitive Factors in Academic Performance. A Critical Literature Review* (Chicago, IL: Consortium on Chicago School Research, 2012), http://eric.ed.gov/?id=ED542543; James J. Heckman, Jora Stixrud, and Sergio Urzua, "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior," *Journal of Labor Economics* 24, no. 3 (2006): 411–82, https://doi.org/10.1086/504455; Jenny Nagaoka, Camille A. Farrington, Melissa Roderick, Elaine Allensworth, Tasha Seneca Keyes, David W. Johnson, and Nicole O. Beechum, "Readiness for College: The Role of Noncognitive Factors and Context," *Voices in Urban Education*, no. 38 (Fall 2013): 45–52; Michelle Richardson, Charles Abraham, and Rod Bond, "Psychological Correlates of University Students' Academic Performance: A Systematic Review and Meta-Analysis," *Psychological Bulletin* 138, no. 2 (2012): 353–87, https://doi.org/10.1037/a0026838

44 Peter N. Stearns, Peter Seixas, and Sam Wineburg, eds, *Knowing, Teaching, and Learning History: National and International Perspectives* (New York: NYU Press, 2000); Sam Wineburg, "Historical Thinking and Other Unnatural Acts," *The Phi Delta Kappan* 80, no. 7 (1999): 488–99.

45 See Heafner and Fitchett, "An Opportunity to Learn U.S. History," for a different discussion of this topic. Note: Heafner and Fitchett used 2010 grade 12 data, which did not include a teacher questionnaire. Moreover, while valid, Heafner and Fitchett took a different analytical approach to defining OTL and have a different set of research questions, making our results difficult to directly compare.

46 Laura S. Hamilton, Julia H. Kaufman, and Lynn Hu, *Preparing Children and Youth for Civic Life in the Era of Truth Decay: Insights from the American Teacher Panel* (Santa Monica, CA: RAND Corporation, 2020), 48, https://www.rand.org/pubs/research_reports/RRA112-6.

47 Hamilton et al., *Preparing Children*.

48 Hamilton et al., *Preparing Children*.

49 Results from the NAEP student questionnaires about course-taking show that 51 percent of students had a class "mainly focused on civics" in the eighth grade, and 72 percent had a class "mainly focused on U.S. history" in the eighth grade. See https://www.nationsreportcard.gov/civics/survey-questionnaires/ and https://www.nationsreportcard.gov/ushistory/survey-questionnaires/.

50 Delroy L. Paulhus and Simine Vazire, "The Self-Report Method," in *Handbook of Research Methods in Personality Psychology,* eds. Richard W. Robins, R. Chris Fraley, and Robert F. Krueger (New York: Guilford, 2007), 224–39, https://www2.psych.ubc.ca/~dpaulhus/research/SDR/downloads/CHAPTERS/2008%20Handbook%20Research%20Methods/paulhus-vazire%2007%20chap.pdf

51 Littenberg-Tobias, "Teaching Citizens."

52 Kahne and Middaugh, *Democracy for Some*.

53 Zuowei Wang, Tenaha O'Reilly, John Sabatini, Kathryn S. McCarthy, and Danielle S. McNamara, "A Tale of Two Tests: The Role of Topic and General Academic Knowledge in Traditional Versus Contemporary Scenario-Based Reading," *Learning and Instruction* 73 (2021): 1–12, https://doi.org/10.1016/j.learninstruc.2021.101462.

54 Peter Levine and Kei Kawashima-Ginsberg, *The Republic Is (Still) at Risk—and Civics Is Part of the Solution* (Medford, MA: Jonathan M. Tisch College of Civic Life, Tufts University, 2017), 3, https://civxnow.org/sites/default/files/resources/SummitWhitePaper.pdf

55 Levine and Kawashima-Ginsberg, *The Republic Is (Still) at Risk*.

56 National Assessment Governing Board, *Civics Framework*, xiv.

57 National Assessment Governing Board, *Civics Framework,* xv.

58 Figure 1 in Hamilton and Parsi, *Monitoring Civic Learning Opportunities*, is based on Michael Hansen, Elizabeth Mann Levesque, Jon Valant, and Diane Quintero, *The 2018 Brown Center Report on American Education* (Washington, DC: Brookings Institution, 2018), https://www.brookings.edu/multi-chapter-report/the-2018-brown-center-report-on-

american-education/, Leveraging Equity and Access in Democratic Engagement, *The LEADE Initiative*, accessed January 30, 2023, https://centerx.gseis.ucla.edu/wp-content/uploads/2018/01/What-Is-Civic-Learning_-LEADE-Initiative.pdf, National Council for the Social Studies, *C3 Framework*, and Raj Vinnakota, From Civic Education to a Civic Learning Ecosystem: A Landscape Analysis and Case for Collaboration (Red & Blue Works, 2019), https://rbw.civic-learning.org/wp-content/uploads/2019/12/CE_online.pdf.

59 Stearns et al., *Knowing, Teaching, and Learning History*, 4.

60 Kadriye Ercikan and Peter Sexias, eds., *New Directions in Assessing Historical Thinking* (New York: Routledge, 2015).

61 Stearns et al., *Knowing, Teaching, and Learning History*; Ercikan and Sexias, *New Directions*; M. Suzanne Donovan and John D. Bransford, eds., *How Students Learn: History, Mathematics, and Science in the Classroom* (Washington, DC: National Academies Press, 2005).

62 College Board, *AP Historic Thinking Skills* (New York: College Board, 2015) https://www.acpsd.net/site/handlers/filedownload.ashx?moduleinstanceid=6536&dataid=25384&FileName=historical_thinking_skills_1_.pdf; "Historical Thinking Skills," American Historical Association, accessed January 23, 2023, https://www.historians.org/teaching-and-learning/teaching-resources-for-historians/teaching-and-learning-in-the-digital-age/the-history-of-the-americas/the-conquest-of-mexico/for-teachers/setting-up-the-project/historical-thinking-skills; National Council for the Social Studies, *C3 Framework*.

63 Deane, P., Song, Y., van Rijn, P. *et al.* The case for scenario-based assessment of written argumentation. *Read Writ* 32, 1575–1606 (2019). https://doi.org/10.1007/s11145-018-9852-7. Sabatini, J., O'Reilly, T., & Deane, P. (2013). Preliminary reading literacy assessment framework: Foundation and rationale for assessment and system design. *ETS Research Report Series*, *2013*(2), i-50; Deane, P., Sabatini, J., Feng, G., Sparks, J.R., Song, Y., Fowles, M., O'Reilly, T., Jueds, K., Krovetz, R., & Foley, C. (2015). *Key practices in the English language arts: Linking learning theory, assessment, and instruction* (Research Report No. RR-15-17). Princeton, NJ: Educational Testing Service. https://doi.org/10.1002/ets2.12063; Deane, P., & Sparks, J.R. (2019). Scenario-based formative assessment of key practices in the English language arts. In H. Andrade, G. Cizek, & R.E. Bennett (Eds.), *Handbook of Formative Assessment in the Disciplines* (pp. 68-96). New York, NY: Routledge; O'Reilly, T., Sabatini, J., Wang, Z., & Dreier, K. (2019). Using scenario-based assessments to measure deep learning. In K. Millis, D. L. Long, J.P. Magliano, & K. Wiemer (Eds.), *Deep Comprehension: Multi-Disciplinary Approaches to Understanding, Enhancing, and Measuring Comprehension* (pp. 197-208). New York: Routledge; Sabatini, J., O'Reilly, T., & Doorey, N. A. (June 2018). *Retooling literacy education for the 21st Century: Key findings of the Reading for Understanding initiative and their implications*. ETS Center for Research on Human Capital and Education. Princeton, NJ: ETS; Sabatini, J., O'Reilly, T., Weeks, J., & Wang, Z. (2019). Engineering a 21st Century reading comprehension assessment system utilizing scenario-based assessment techniques. *International Journal of Testing 20*(1), 1-23. https://doi.org/10.1080/15305058.2018.1551224.

64 Paul Deane and Jesse Sparks, "Scenario-Based Formative Assessment of Key Practices in the English Language Arts," in *Handbook of Formative Assessment in the Disciplines*, eds., Heidi L. Andrade, Randy E. Bennett, and Gregory J. Cizek (New York: Routledge, 2019), 68–96.

65 Blair Lehman, G. Tanner Jackson, and Carol Forsyth, "A (Mis) Match Analysis: Examining the Alignment between Test Taker Performance in Conventional and Game-Based Assessments," *Journal of Applied Testing Technology* 20, no. S1 (2019): 17–34, http://jattjournal.net/index.php/atp/article/viewFile/142699/101509; Steven L. Wise, "Rapid - Guessing Behavior: Its Identification, Interpretation, and Implications," *Educational Measurement: Issues and Practice* 36, no. 4 (2017): 52–61, https://doi.org/10.1111/emip.12165.

66 Yoav Bergner and Alina A. von Davier, "Process Data in NAEP: Past, Present, and Future," *Journal of Educational and Behavioral Statistics* 44, no. 6 (2018), 706–32, https://doi.org/10.3102/1076998618784700; Madeleine Keehner, Joanna S. Gorin, Gary Feng, and Irvin R. Katz, "Developing and Validating Cognitive Models in Assessment," in *The Wiley Handbook of Cognition and Assessment: Frameworks, Methodologies, and Applications,* eds. Andre A. Rupp and Jacqueline P. Leighton (New York: Wiley Blackwell), 75–101.

67 For example, see the Delaware System of Student Assessments, Social Studies Training Test, Grade 4 Training Test (General Form), http://delaware.pearsonaccessnext.com/tutorial/ss-item-samplers/, sample question 6.

68 See https://enaep.cotw-ng.naepnpd.org/ade/index.html?mode=cotw for a released reading question from the 2022 NAEP reading assessment that uses "select-in-text" functionality.

69 See https://enaep.cotw-ng.naepnpd.org/ade/index.html?mode=cotwfor a released reading question from the 2022 NAEP reading assessment that uses "select-in-text" functionality.

70 Julie Coiro, Carla Coscarelli, Cheryl Maykel, and Elena Forzani, "Investigating Criteria That Seventh Graders Use to Evaluate the Quality of Online Information," *Journal of Adolescent & Adult Literacy* 59, no. 3 (2015): 287–97, https://doi.org/10.1002/jaal.448; Jesse R. Sparks and Paul Deane, *Cognitively Based Assessment of Research and Inquiry Skills: Defining a Key Practice in the English Language Arts,* ETS Research Report No. RR-15-35 (Princeton, NJ: ETS, 2015), https://doi.org/10.1002/ets2.12082

71 Randy Elliot Bennett, Hilary Persky, Andrew R. Weiss, and Frank Jenkins, *Problem Solving in Technology-Rich Environments. A Report from the NAEP Technology-Based Assessment Project, Research and Development Series*, NCES 2007-466 (Washington, DC: National Center for Education Statistics, 2007); Ion Madrazo Azpiazu, Nevena Dragovic, Maria Soledad Pera, and Jerry Alan Fails, "Online Searching and Learning: *YUM* and Other Search Tools for Children and Teachers," *Information Retrieval Journal* 20, no. 5 (2017): 524–45, https://doi.org/10.1007/s10791-017-9310-1; Els Kuiper, Monique Volman, and Jan Terwel, "The Web as an Information Resource in K-12 Education: Strategies for Supporting Students in Searching and Processing Information," *Review of Educational Research* 75, no. 3 (2005): 285–328, https://doi.org/10.3102/00346543075003285; Amber Walraven, Saskia Brand-Gruwel, and Henry P. A. Boshuizen, "Information-Problem Solving: A Review of Problems Students Encounter and Instructional Solutions," *Computers in Human Behavior* 24, no. 3 (2008): 623–48, https://doi.org/10.1016/j.chb.2007.01.030.

72 This constrained design is consistent with simulated search tools developed in previous assessments. Bennett et al., *Problem Solving*; Julie Coiro, Jesse R. Sparks, and Joanna M. Kulikowich, "Assessing Online Collaborative Inquiry and Social Deliberation Skills as Learners Navigate Multiple Sources and Perspectives," in *Handbook of Multiple Source Use,* eds. Jason L. G. Braasch, Ivar Bråten, and Matthew T. McCrudden (New York: Routledge, 2018), 485–501; Donald J. Leu, Julie Coiro, Joanna Kulikowich, and Weiwei Cui, "Using the Psychometric Characteristics of Multiple-Choice, Open Internet, and Closed (Simulated) Internet Formats to Refine the Development of Online Research and Comprehension Assessments in Science: Year Three of the ORCA Project" (paper presented at the 62nd annual meeting of the Literacy Research Association, San Diego, CA, December 2012).

73 For one example of a digital map that explores change over time, see "Justice Deferred: Executive Order 9066 and the Geography of Japanese American Imprisonment," esri, accessed January 29, 2023, https://storymaps.esri.com/stories/2017/japanese-internment/index.html

74 "NAEP Technology & Engineering Literacy (TEL) Report Card: Chicago Task," The Nation's Report Card, accessed January 29, 2023, https://www.nationsreportcard.gov/tel/tasks/chicago/

75 Danielle Picard and Derek Bruff, "Digital Timelines," Vanderbilt University, accessed January 29, 2023, https://cft.vanderbilt.edu/guides-sub-pages/digital-timelines/.

76 *National Education Statistics Act*, 20 *USC* 9001 (2001), §§ 1 et seq.

77 Ina V. S. Mullis, "White Paper on 50 Years of NAEP Use: Where NAEP Has Been and Where It Should Go Next" (NAEP Validity Studies [NVS] Panel) (American Institutes for Research, San Mateo, CA, 2019), https://www.air.org/sites/default/files/2021-06/50-Years-of-NAEP-Use-June-2019.pdf.

78 Kajsa Yang Hansen and Rolf Strietholt, "Does Schooling Actually Perpetuate Educational Inequality in Mathematics Performance? A Validity Question on the Measures of Opportunity to Learn in PISA," *ZDM* 50, no. 4 (2018): 643–58, https://doi.org/10.1007/s11858-018-0935-3.

79 Educating for American Democracy, *Roadmap to Educating for American Democracy* (Cambridge, MA: iCivics, 2021), 3, https://www.educatingforamericandemocracy.org/wp-content/uploads/2021/02/Roadmap-to-Educating-for-American-Democracy.pdf

80 Black, "Constitutional Compromise," 741.

81 For more information about the NAEP sampling procedure, please see National Center for Education Statistics, *National Assessment of Educational Progress (NAEP) 2018 Civics, Geography, and U.S. History Grade 8 Assessments Restricted-Use Data Files Data Companion,* NCES 2021050 (Washington, DC: U.S. Department of Education. Institute of Education Sciences, 2021).

82 All surveys can be accessed here: https://nces.ed.gov/nationsreportcard/experience/survey_questionnaires.aspx.

83 The AM statistical software program, developed at the American Institutes for Research (AIR), can be found here: https://am.air.org/.