

NATIONAL

HISTORY DAY EVALUATION

FINAL REPORT

November 2010

Kay Sloan
Saul Rockman

Rockman et al
Independent · Insightful · Informative

Acknowledgements

Many people who make National History Day a rewarding experience for thousands of students also made this study possible. We especially wish to thank Cathy Gorn, Ann Claunch, and Lynn Fontana, for inviting us to undertake this study and guiding our work each step of the way. We also wish to thank the NHD state and site coordinators, Becky Butz, Steve Cure, Mary Catherine Marshal, Nancy Norris-Bauer, and Margaret Wrenn, who paved the way for us to gather data in the four study sites, and our other key site contacts, Ann Carlock, Gail Ingram, and Deborah Johnson, who helped us find comparison classes, schedule meetings, administer surveys, assemble assessment data—and see the depth and range of History Day implementations. Many thanks also go to the teachers and students who took part in surveys and interviews, shared their stories, and helped us understand how much heart and hard work go into a National History Day project. Finally, we want to express our gratitude to Kenneth E. Behring, for his support of this study and the National History Day program.

We also thank Candy Miller for her skill in analyzing the extensive data collected for the study, and Katie White Walters and Justin Robertson for their help in conducting the study and compiling the report.

Kay Sloan

Saul Rockman

Table of Contents

Introduction	1
Research Design	2
Recruitment and Study Sites	2
Instruments and Data Collection	4
Comparison Groups, Sample Sizes, and Demographics	8
Data Analysis and Scoring	11
Parameters of the Study	11
<i>Key Findings</i>	13
<i>Discussion of Findings</i>	17
Applied Skills	17
Confidence in Applied Skills	18
Application of Skills	19
Writing Performance	22
Other 21 st Century Skills	
Academic Performance	24
Social Studies Performance	26
Performance in Other Academic Areas	27
Interest and Engagement	31
Interest in History	31
Civic Engagement	34
Differential and Cumulative Impact	36
Impact on Different Kinds of Students	36
Impact over Time	40
Conclusions and Recommendations for Further Study	46
Appendix A: Sample Data Request	49
Appendix B: Scoring Rubric for Student Writing	51

Index of Tables

Table 1. Numbers of Middle and High School Respondents, Surveys and Performance Assessments	8
Table 2. Breakdowns by Gender, by Site, Surveys and Performance Assessments	9
Table 3. Ethnicity by Site, Surveys and Performance Assessments	10
Table 4. Ethnic Distribution Compared to U.S. Public Elementary and Secondary School Enrollment	10
Table 5. Composite Scores on Performance Assessments	21
Table 6. High School Performance Assessment by Grade	21
Table 7. Percentages of Students Receiving Scores of 1-6 on Pre- and Post- Writing Assessments	22
Table 8. Writing Performance Mean Scores, by State	24
Table 9. NHD and Comparison Groups' Mean Scores, PASS test in Social Studies, 8 th Grade	28
Table 10. NHD and Comparison Groups' Performance Level, PASS test in Social Studies, 8 th Grade	28
Table 11. NJ ASK Language Scores prior to and after NHD Participation	29
Table 12. Composite Means for Pre- and Post-Survey Items on Civic Engagement	34
Table 13. High School Performance Assessment By Gender and Ethnicity	37
Table 14. Writing Performance by Gender and Ethnicity	38
Table 15. Differences between NHD and Comparison Group Students Based on Gender and Levels of Interest, Engagement, and Confidence	39

Index of Figures

Figure 1. Students' Confidence in Applied Skills	18
Figure 2. Writing Scores, Pre- and Post-Test	23
Figure 3. Percentages of NHD and Comparison-Group Students with High and Low Writing Scores	23
Figure 4. Students' Interest in History Classes	31
Figure 5. Students' Interest in History Classes Compared to Other Classes	31
Figure 6. Students' Interest in Historical Topics	32
Figure 7. Students' Engagement in Current & Civic Issues & Events	35
Figure 8. Correlations between Impact and Interest, Engagement, and Confidence	39
Figure 9. Students' Performance Assessment Composite Scores, by Years of NHD Participation	41
Figure 10. Grades Based on Years of NHD Participation	42
Figure 11. Percentages of Commended Performance on TAKS Based on Years of NHD Participation	43
Figure 12. Commended Performance Based on Years of Participation	44
Figure 13. Composite Means by Years of Participation	45

HISTORY DAY EVALUATION

FINAL REPORT

Introduction

In 2006 the National History Day organization asked Rockman et al to develop a research plan to explore the impact of the program, which, over three decades, had grown from a small day-long contest, first held in Cleveland, Ohio, in 1974, to a national program attracting more than half a million students annually from every state in the union, Washington, D.C., American Samoa, Guam, international schools, and Department of Defense schools in Europe. District and state contests still take place on a single day, but the Kenneth E. Behring national contest, held in June at the University of Maryland in College Park, now lasts almost a week.

During that time, students in grades 6-12 who have risen through the ranks of district and state contests present papers, exhibits, performances, documentaries, and websites to a panel of judges. Students select topics based on personal interests and an annual theme, which changes each year. The 2010 theme, “Innovation in History: Impact and Change,” inspired presentations on everything from Galileo to nylon hose to *Sesame Street*. The range was evident even among students from the four states participating in this study: “From Bayous to Beachheads” about Higgins boats, used in the Normandy landing (Texas); “Paterson: Industry Began Here!” about America’s earliest manufacturing, powered by the falls of the Passaic River (New Jersey); “The Blues Had a Baby and They Called It Rock and Roll: Leonard Chess and the Integration of Pop” (Colorado); and “MASH: An Innovation in Battlefield Medicine” (South Carolina).

When not rehearsing or presenting their work, students, families, and teachers meet with Congressional representatives, tour the National Archives and other landmarks in the nation’s capital, talk to *HISTORY* interviewers, get to know other contestants, and, new for 2010, follow a live NHD Twitter feed. As the NHD organization says, “It’s not just a day, it’s an experience.”

Even students who do not enter or make it through the contests put in many hours on their NHD projects. In groups or individually, most work throughout the school year, during and after school, gathering primary and secondary sources from school and university libraries, archives, museums, oral history interviews, and historic sites, refining their topic, defining its historical context and significance—and mustering the skills and confidence for juried contests. Calling the experience “history year” would not be far off.

As NHD has grown so have the numbers of testimonials from students, teachers, and parents crediting the program with helping students develop vital research, critical thinking, and communication skills. Praise and gratitude come not just from school-age participants, but also from NHD alumni who have gone on to careers in media, marketing, law, medicine, education, and other fields where, they say, skills and knowledge acquired through NHD have served them well.

What was absent from the rich store of testimonials was independently gathered empirical data—the hard, evidentiary proof about program effectiveness that school administrators need to select course offerings and allot staff and funds. As budgets shrink and accountability pressures rise, they more and more need answers to critical questions: does NHD affect students’ performance on high-stakes tests? What specific skills do students gain? Who benefits from participation? Does NHD help teachers meet standards?

To explore those questions, Rockman designed a study to examine students’ performance on state standardized assessments, looking not just at social studies but also at other academic subjects where students’ skills might transfer. The study also included performance assessments designed to measure students’ ability to apply the research, writing, and critical thinking skills developed through NHD participation—skills that track closely with the 21st Century skills identified by educators and business leaders as the skills students need to enter college and the workplace fully prepared.

The national discourse about what young people should learn in school also includes conversations about the need to impart knowledge about the history and culture we share. A recent report from Common Core, entitled “Still at Risk: What Students Don’t Know, Even Now,” indicates that too many students can’t name the German Chancellor during World War II or place the Civil War in the correct half-century—a lack that may put not just students and the workforce, but also the nation at risk. The new Common Core Standards, launched by the National Governors Association and Council of Chief State Schools Officers, emphasize, among other skills, the importance of content knowledge and include historical documents in suggested reading.² NHD is at its core a history project, and, in addition to looking at students’ performance on statewide social studies assessments, the NHD study included survey questions designed to gauge students’ interest in past and current events and their understanding of historical context and narrative.

Research Design

Recruitment and Study Sites

In Fall 2008, with funding from NHD and Kenneth E. Behring, Rockman began recruiting sites for the study. Our criteria included geographic representation, diversity in the student population and inclusion of underrepresented minorities, sufficient history with NHD to allow researchers to look at student performance over time, and implementations that reflected fidelity to national program requirements.

¹ Frederick M. Hess, “Still at Risk: What Students Don’t Know, even Now: A Report from Common Core” (Washington, D.C.: Common Core, 2008), 7-9.

² The Common Core standards and background information on their development are available at <http://www.corestandards.org/>.

Participation in the study also required instructional time at the beginning and end of the school year, the cooperation of the district assessment offices in providing assessment and other institutional data, and a data system robust enough for local staff to retrieve data electronically or without undue effort.

Delays in recruiting and securing permission to conduct the research led to a postponement of the full study until the 2009-2010 school year—but also, fortuitously, allowed us to test instruments and plans in a Spring 2009 pilot in the Aldine Independent School District in Houston, Texas. Interviews with the Aldine social studies coordinator and middle and high school teachers gave us valuable background information. Data from over 2,000 high school students, and from matched comparison groups of around 200 NHD and non-NHD students, showed that NHD participants were outperforming their peers on multiple measures and indicated that upward trends in performance were linked to multiple years of participation.

The 2009-2010 research took a closer look at these gains and trends among Texas students, and expanded the study to three other school districts that provided urban, suburban, and rural settings, and the range we sought in implementation and student populations (see pp. 11-12 below for student demographics). The four sites participating in the national evaluation were:

- **Aldine Unified School District**, Houston, Texas. AUSD is a large, diverse urban district north of Houston that enrolls some 60,000 students in 70 schools. Eighty percent of those students are considered disadvantaged. For its efforts to meet students' needs, Aldine has been a nominee and winner of the Broad prize, which honors districts serving high-need urban populations and reducing achievement gaps. Aldine also has a long history with NHD, and is known around the state as a formidable competitor—a status recognized locally by the fact that students can letter in NHD. The program is implemented in middle and high schools, in regular education and gifted classes, in social studies and media classes, and in International Baccalaureate “Theory of Knowledge” classes.
- **Paterson School District**, Paterson, New Jersey. Also an urban site, the Paterson school district is the third largest school district in New Jersey, with an enrollment of 24,000 students in 52 schools. The student population is highly diverse, and includes students of Hispanic, African-American, Middle Eastern, Asian, and Caucasian descent. Nearly 50 percent of all students in Paterson speak a primary language other than English, with a total of 37 languages spoken in district schools. Though the Paterson schools face challenges stemming from budget cuts and economic downturns, they have strong NHD programs for students at multiple grade levels, their teachers have won state NHD awards, and they routinely send students to the national contest.
- **Chesterfield County Schools**, Cheraw and Chesterfield, South Carolina. The Chesterfield County Schools, in the north central part of the state, include 16 schools and a total enrollment of 8,000 students, 57 percent of whom are eligible for free or reduced lunch. In 2009, this small, rural district led the state in student performance on the new end-of-course test in U.S. History. At Cheraw High School, ninth graders in Honors World History participate in NHD as a class requirement; most also

participated in eighth grade, where NHD was mandatory for all students until 2009-2010. NHD students compete at state contests and many continue, often successfully, at the national finals, in what their teacher calls the “academic Olympics.”

- **A large urban/suburban district in Colorado.**³ The fourth site, the study’s western state site, enrolls a diverse population of 78,000 students in 161 schools. In addition to implementing the program in regular education classes, the district includes NHD in its school of the arts curriculum and as part of its International Baccalaureate program in European History. Like other sites, the site includes teachers who have a long history with NHD, and a long list of students who have won awards at state and national contests, including students who have repeatedly won awards for their documentary, multimedia films and for their group and individual dramatic performances.

Instruments and Data Collection

The instruments and data collection for the study, described in detail below, were designed to explore key questions about the impact of NHD participation, which also frame the report:

- *What skills do students gain from NHD participation, and, compared to their peers, how successfully can they apply them?*
- *Does NHD have a positive effect on students’ performance on high-stakes tests —not just in social studies but also in other academic subjects?*
- *How do NHD students’ interests in history, and their perspective on past and current events, compare to their peers’?*
- *Does NHD have a positive impact on all students, and does impact build over time?*

Performance Assessments. The existing anecdotal evidence, along with student and teacher focus groups and surveys of NHD students and alumni, both conducted in preparation for this study, pointed to a set of research and critical thinking skills that students develop through participation. To see how well students could apply these skills in other contexts, and how NHD students compared to non-participating counterparts, we developed performance assessments with multiple-choice and short-answer questions that asked students to identify primary sources, explain how they would select and evaluate sources, and interpret or draw conclusions from historical information.⁴ Short-answer questions were based on reviews of state social studies assessments, NAEP history assessments, reviews of the literature, and consultations with history teachers; some items were taken directly from publicly available NAEP items, and from the Texas Assessment of Knowledge and Skills (TAKS).⁵ Permission to use the TAKS items came from TAKS and Pearson Publishing.

³ The school district requested that study reports not identify it by name.

⁴ R. J. Marzano, D. Pickering, and J. McTighe, *Assessing Student Outcomes: Performance Assessment Using the Dimensions of Learning Model* (Alexandria, VA: (Association for Supervision and Curriculum Development, 1993), 123.

⁵ See, for NAEP, <http://nces.ed.gov/nationsreportcard/about/booklets.asp>; for TAKS, http://www.tea.state.tx.us/index3.aspx?id=3839&menu_id=793.

The assessments also asked students to interpret historical images or political cartoons, and write an essay arguing for or against an issue (pre-test) or explaining their position on an issue depicted in a cartoon (post-test). Most images or cartoons were in the public domain; Rockman secured appropriate permissions to use those from more recent sources.

Political or editorial cartoons often portray or provoke some controversy, and the biggest challenge in developing assessments was finding images or cartoons that were accessible and recognizable, but not so controversial as to elicit inappropriate responses. More interested in gauging students' critical thinking skills than specific historical knowledge, and mindful of the fact that students' shared knowledge might be limited, we searched for familiar figures and events. Interestingly, images initially considered to be good candidates ran the risk of eliciting inappropriate responses when those images began appearing on August 2009 town-hall protest signs. One cartoon under consideration, for example, entitled "Filling Station," showed Hitler filling up the heads of young men.⁶ In our pilot, students pointed out that students who had seen signs portraying the president as Hitler or accused of brainwashing might reference that knowledge in their responses. While in a regular classroom that could create a teachable moment, we did not want to tap into negative opinions that might affect performance or scoring or create difficulties for participating teachers.

The final fall, pre-assessments included, for middle schoolers, a painting depicting the first Thanksgiving in rich detail. Citing those details to support their interpretations, students were asked to choose among different statements about the artist's message: Did the painting show, for example, that the Pilgrims couldn't have survived without the Native Americans, that Native Americans were starving until the Pilgrims arrived, or that the Pilgrims welcomed Native Americans to their bountiful feast. High school students were asked to explain the symbols and actions in two cartoons about women's suffrage: one showed a donkey and elephant intent on fishing and ignoring a drowning female—until she was holding a fish labeled "vote"; another showed both a donkey and elephant vying for female attention.⁷ For the spring, post-assessment, the cartoons were the same for middle and high school students, and we combined the interpretation of editorial cartoons with the writing assessment.

The writing topics came from existing assessments and consultation with teachers. Again the goal was not so much to test students' historical knowledge as to gauge their ability to frame, support, and sustain an argument—skills they might gain through or need for history classes and projects. To ensure that students could complete the writing assessment in the time allotted (one class period), we chose topics on which students would likely have an opinion, and gave students choices. Options on the fall or pre-assessment were whether or not the legal driving age should be raised to eighteen, or whether U.S. History or World

⁶ The Hitler cartoon, from the British Cartoon archive housed at the University of Kent, is available at: [http://www.cartoons.ac.uk/browse/cartoon_item/anytext=world%20war%20ii?subjects_text\[\]=World%20War%201939-1945&personalities_text\[\]=Hitler;%20Adolf%20\(1889-1945\)%20\[Hit\]&page=138](http://www.cartoons.ac.uk/browse/cartoon_item/anytext=world%20war%20ii?subjects_text[]=World%20War%201939-1945&personalities_text[]=Hitler;%20Adolf%20(1889-1945)%20[Hit]&page=138).

⁷ *The First Thanksgiving 1621*, by J.L.G. Ferris, came from the Library of Congress, <http://www.loc.gov/teachers/classroommaterials>.

History should be taught in eighth grade.⁸ The spring assessment included three editorial cartoons about the impact of technology: one depicted cavemen decrying the use of fire as the end of civilization; a second showed a character whose absorption in cyberspace drew his attention away from everyday activities; in the third cartoon, a solitary young person tried to make friends through social networking while real friends stood outside the door. Students had to describe what all three cartoons depicted, but could opt to write about a single cartoon or all three in their essays. Scoring rubrics for the writing were based on nationally available and widely used rubrics. (See p. 22 for further discussion of the rubric and scoring, and Appendix B, p. 53, for the rubric used for both the pre- and post-test of writing.)

Achievement, Demographic, and Behavioral Data. Rockman reviewed annual standardized assessments administered in each state, then submitted data requests to each site for students' scores on appropriate assessments. We asked for current and historical data (2009-2010 and two-three years prior) for all students, in all grades in which students participate in NHD. In South Carolina, we also requested assessment data for students from the middle school where, until this year, NHD was mandatory for all 8th graders, and from a comparable middle school in the same district.

We also requested GPA's, grades, end-of-course assessment scores; and other institutional data such as attendance rates, suspensions or behavioral referrals, and graduation rates. These data allowed Rockman to draw a profile of each district, select comparison groups that as closely as possible matched the NHD group on demographics and prior achievement, and track the academic performance of students over multiple assessments and multiple years. (See Appendix A, pp. 51-52, for a sample data request.)

Because states do not administer the same standardized tests, and courses and grading policies are locally determined, the findings about NHD's impact on students' academic performance are reported by site, and not in aggregate, or across states. Although we requested data from national assessments, such as Advanced Placement (AP), International Baccalaureate, and the National Assessment of Educational Progress (NAEP), and report some findings about NHD students' performance compared to that of other students in their school district, we were not able to conduct meaningful analyses across sites, either because sample sizes were too small, sites did not archive scores, or, in the case of NAEP, tests were not administered with enough regularity. (See pp. 28-29 for further discussion of state assessments.)

Student Surveys. Surveys included items tested during background and pilot research activities. In addition to demographic and background information, surveys asked students about their interest in history and history classes; their confidence in research, communication, and study skills; and the impact of NHD participation (NHD-only). Although dispositions may develop in more gradual and subtle ways, NHD can also help create an informed, engaged citizenry. To see how NHD students' interest in current as well as past events and their sense of civic responsibility and engagement differed from their peers', we also included, in the survey, items adapted from existing national surveys designed to gauge these

⁸ Cartoons about women's suffrage, from the Opper Project, are available at <http://hti.osu.edu/opper>.

perceptions.⁹ Items focused on the extent to which students seek out online and other news sources to stay informed, whether they engage in discussions about local or national issues, and whether they feel they have the power and responsibility to make a difference.

Teacher Surveys. We asked NHD and non-NHD teachers about the courses they teach and the frequency with which students in those classes engage in long-term projects requiring research. We also asked NHD teachers about their experience with and role in the program, and their perceptions of the program's impact on students—including students with unexpected engagement and success in the program.

Interview Protocols. Interview questions for NHD teachers and students were based on previous focus group questions, survey responses, and NHD activities unique to each site. Student questions focused on students' reasons for participating, the topics and projects they chose, and the skills they gained from researching topics, working in groups, presenting before real audiences, and fielding judges' questions. Questions for teachers mirrored these items; researchers also gathered additional information on teachers' history with the NHD program, classroom and school implementations, and perceptions of students' experiences.

The NHD organization reviewed the surveys, which were piloted in spring 2009. NHD, history teachers, and a state coordinator reviewed the performance assessments. Rockman conducted a pilot of the assessments with 25 high school students and their teacher, during which students completed the assessments, then provided feedback on readability, clarity, and appropriateness; the teacher timed students, documented comments, and shared her own feedback.

The evaluation team collected performance assessment and survey data from NHD and comparison-group students twice during the 2009-2010 school year, once in fall 2009 and again in spring 2010. Researchers visited each site in spring 2010, and teachers completed post-only surveys. All instruments were made available on paper and online. Rockman submitted requests for standardized test data, grades and GPAs, and other institutional data early in the second semester of the school year, and reviewed requests with district research and assessment personnel during site visits. Rockman received human subjects approval for the study from an independent institutional review board, and, in the case of the Colorado district, also received approval from the district's board.

All instruments are available upon request.

⁹Items adapted from the Tufts National Survey of Civic Participation, available at <http://activecitizen.tufts.edu/downloads/questionnaire.pdf>; and the UC-Berkeley Service-Learning Research and Development Center's survey of civic responsibility, available at http://www.peecworks.org/PEEC/PEEC_Inst/0038CEB0-007EA7AB.1/furcosurvey%20civicrespons.pdf.

Comparison Groups, Sample Sizes, and Demographics

The study design called for comparison groups, to enable researchers to determine how NHD students' academic performance, applied skills, and interests differed from their peers'. For the primary data—the surveys and performance assessments—we requested, wherever possible, an internal comparison group of students with similar ability levels, in a similar class within the same school. To ensure that we were making fair comparisons, we also asked that teachers in comparison-group classes engage students in project-based learning that resembled the kind and level of work required of NHD students. Where there were no comparable classes within a school, we used the same criteria, but looked to other schools in the same district with a demographic profile similar to the NHD school, and with social studies teachers whose projects required research, writing, long-term assignments, and self-directed learning.

In most cases, comparison-group samples far exceeded our sample of NHD students. Rather than comparing NHD students to all other students, and running the risk of comparing apples to oranges or privileging a certain kind of student, we created matched comparison samples, based on a beginning test score; these varied somewhat depending on state assessments, but were typically language arts scores. We then selected a specific number of students based on how they fell within each percentile rank on the assessment, while ensuring that the comparison group had an equivalent number of students based on ethnicity and gender.

Although we started with data from close to 7,000 students, the final matched samples for secondary data included approximately 1,500 students. The final sample for the primary data (surveys and performance assessments) included 48 middle school students in Texas and Colorado (neither South Carolina nor New Jersey included middle schools), and 410 high school students, for a total student sample of 458 students, 274 of whom were NHD students, and 184, comparison-group students. (See Table 1.)

Because of some attrition, either on the part of students who, during the school year, opted out of NHD, or by teachers whose schedules or school obligations left too little time for NHD, our numbers for students completing surveys and performance assessments are lower than expected, but large enough to conduct various analyses. This is especially true of the writing assessment, which could typically involve a smaller sample of students. Numbers varied somewhat by instrument and pre- and post- responses; tables, figures, and discussions include N's of NHD and comparison-group students.

Table 1. Numbers of Middle and High School Respondents, Surveys and Performance Assessments

Site	Middle School			High School		
	NHD	Comparison	TOTAL	NHD	Comparison	TOTAL
Pre-	35	28	63	281	278	559
Post-	50	26	76	184	251	435
Both	33	15	48	241	169	410
TOTAL for Surveys and Performance Assessments (48 + 410)						458

Demographics for Primary Data (Surveys and Performance Assessments)¹⁰

School breakdowns by gender were fairly similar and mirrored district figures in Texas and South Carolina. In New Jersey, there were more males than females in the NHD group, and more females in the comparison group. In Colorado, males outnumbered females in both groups. (See Table 2.)

Table 2. Breakdowns by Gender, by Site, Surveys and Performance Assessments

Site	NHD	Comparison
TX (N=170)		
Male	57%	55%
Female	44%	45%
SC (N=132)		
Male	59%	51%
Female	41%	49%
NJ (N=59)		
Male	58%	39%
Female	42%	62%
CO (N=93)		
Male	73%	63%
Female	27%	37%

Source: Student Surveys

Ethnic distribution in some cases mirrored district and state distributions. In Texas, there were higher percentages of Asian students in the study than in the district or state; percentages of White students in the NHD population were more similar to state than district numbers; comparison group percentages reflected district-wide percentages. In New Jersey, there were more NHD Asian students than in the comparison group, district, or state; the representation of Black students in the NHD group was closer to state percentages; in the comparison group, representation was closer to district figures. In Colorado, percentages of White students in NHD and comparison groups was similar to state figures, but higher than district figures; percentages of Hispanic students was lower in NHD and comparison groups than in the district or state. (See Table 3; 449 of the 458 students answered the race/ethnicity question.)

Compared to U.S. public school enrollment figures, representation of Black and Hispanic students was somewhat higher in the study sample than in the population as a whole—confirming that we achieved a study goal of oversampling under-represented populations—and the representation of White students, somewhat lower. (See Table 4.)

¹⁰ The demographics reported here are based on student survey responses from NHD and non-NHD students. We can assume that the demographics for the secondary data are very similar because, in creating comparison groups, we matched based on ethnicity and gender as well as prior achievement.

Table 3. Ethnicity by Site, Surveys and Performance Assessments

	Comparison	NHD	District	State
Texas (N=165)				
American Indian/Alaska Native	3%	3%	<1%	<1%
Asian/Pacific Islander	12%	12%	2%	3%
Black/African-American	29%	23%	31%	14%
Hispanic/Latino(a)	55%	56%	64%	47%
White/Caucasian	7%	20%	4%	35%
Other	4%	1%		
South Carolina (N=132)				
American Indian/Alaska Native	3%	0%	1%	<1%
Asian/Pacific Islander	2%	1%	1%	
Black/African-American	27%	32%	39%	39%
Hispanic/Latino(a)	2%	3%	3%	5%
White/Caucasian	73%	55%	55%	53%
Other	2%	7%	2%	
New Jersey (N=59)				
American Indian/Alaska Native	0%	0%	<1%	<1%
Asian/Pacific Islander	9%	23%	4%	8%
Black/African-American	39%	15%	41%	17%
Hispanic/Latino(a)	49%	50%	53%	19%
White/Caucasian	3%	4%	4%	55%
Other	15%	8%		
Colorado (N=93)				
American Indian/Alaska Native	0%	1%	1%	1%
Asian/ Pacific Islander	0%	6%	4%	3%
Black/African-American	32%	17%	16%	6%
Hispanic/Latino(a)	9%	15%	54%	28%
White/Caucasian	55%	64%	25%	61%
Other	15%	8%		

Source: Student surveys and district and state departments of education websites

Table 4. Ethnic Distribution Compared to U.S. Public Elementary and Secondary School Enrollment

	Comparison	NHD	U.S.
American Indian/Alaska Native	2%	1%	1%
Asian/ Pacific Islander	<1%	<1%	5%
Black/African-American	30%	23%	17%
Hispanic/Latino(a)	31%	30%	21%
White/Caucasian	34%	39%	56%
Other	2%	5%	

Source: National Center for Educational Statistics¹¹

¹¹ See http://nces.ed.gov/programs/digest/d09/tables/dt09_041.asp.

Data Analysis and Scoring

We used varied methods to analyze the school data received for NHD and comparison-group students. First, we used descriptive and frequency analysis to look at grades, discipline referrals, and attendance data, where available, and performance on state standardized tests based on performance level in various subjects, typically social studies, language arts, math, and science. In addition, we used paired sample and independent t-tests of means to examine year-to-year change and differences between groups. ANOVA and regression analysis was also used to look at differences in test scores based on student demographic variables, such as gender, ethnicity, gifted status, free or reduced lunch status, and number of years of participation in the program.

For most survey and performance assessment items, we analyzed basic frequencies and descriptives and ran cross-tabs to examine differences based on students' years of participation in NHD, gender, race or ethnicity, and site. Both the pre- and post- student surveys included identical sets of scaled items about students' 1) interest in historical periods, themes, or issues; 2) confidence in research, writing, and presentation skills); and 3) engagement in current events and issues. We compared pre-survey and post-survey responses from all NHD students (N=272) to those from all non-NHD students (N=183), looking at means for both groups, and conducting t-tests to examine between-group differences and calculate statistical significance.

We also created composite interest, engagement, and confidence mean scores for both sets of students, looking at differences between groups on both the pre- and post- surveys as well as pre- to post- changes, and running tests for significance. Using these three composite scores, we also looked at differences by state and by gender. The post-survey also included an item for NHD students about the perceived impact of NHD participation. Using regression analysis, we looked at the relationship between NHD students' perceptions of impact and their self-reported levels of interest, engagement, and confidence.

In analyzing open-ended responses, we looked through all responses for recurrent themes. We then selected what seemed to be representative responses, then pulled every tenth response to see how those compared. We reviewed interview transcripts for more in-depth analyses and quotes.

For the writing assessments, we developed a scoring rubric based on the NAEP persuasive essay rubric and the 6+1 Writing Traits rubric. (See p. 21 for rubric citations.) Scorers were trained using the rubric and benchmark essays. Three researchers scored a sample of essays from each site, with site identifiers removed. At intervals during the scoring process, a second reader scored randomly sampled essays to ensure consistency and inter-rater reliability.

Parameters of the Study

The attrition noted above and non-responses reduced the power and effect sizes for the study to a certain degree. They could also have introduced some selection bias, because NHD may attract a certain kind of student and requires sustained hard work. Middle school participation in the study was low, in part

because two sites did not include middle schools, and a third was not able to recruit a comparison middle school. We also lacked a comparison high school in one site.

Students are also exposed to NHD to different degrees: some schools and teachers make NHD the main focus of a special class, some incorporate the program into an existing class, some conduct an after-school program, and some do a combination of these. While the variety in implementation makes NHD a richer program, and gives teachers and schools more latitude, it makes it very difficult to control the treatment in a research study.

We have tried to offset imbalances, biases, or implementation differences with multiple instruments, triangulation, and confirmations of findings in both primary and secondary data. We have taken sample sizes into account when conducting tests for significance, and where possible also included more middle and comparison group students in the analysis of secondary data. Discussions of findings indicate numbers of students included in each analysis, and where possible describe other factors that might account for differences.

This study cast a wide net to explore the impact of NHD and compare participants to their peers. We mined results from a variety of state tests; assessed research, writing, and other skills students might acquire through participation; and gathered data on students' interest in past and current events and their perceptions of civic engagement. So broad a study could not examine NHD's impact on specific student populations or link impact to specific classroom implementations. The Conclusions and Recommendations section (see pp. 48-50) outlines further, more controlled studies that could investigate impact, in context, in detail, and over time.

Applied Skills

(Data sources: Performance Assessments¹², N=458; focus groups and interviews, N=75)

- **NHD students already know how to do college-level research.** Performance assessments showed that NHD students' ability to find, evaluate, and use information exceeds their peers'. By a margin of 2 to 1, NHD students correctly identified primary sources. Differences between middle-school students were smaller, but still evident.

NHD students understand the value of multiple sources and know to look beyond basic sources listed by non-NHD students—books, newspapers, textbooks, and Wikipedia—to experts, personal contacts, museums, lecture notes, diaries, journals, films, first-person accounts, and biographies. NHD students also have a more sophisticated understanding of how to evaluate sources. They can tell a good source from a bad one. While peers suggested checking publication date or author, NHD students said they would check for a valid copyright and reputable publisher, look for .edu rather than .com sources, confirm authors' credibility by cross-checking other references, and look for corroboration across sources.

- **NHD students are critical thinkers who can digest, analyze, and synthesize information.** Performance assessments show that, overall, NHD students were significantly ($p < .001$) better than their peers at interpreting historical information, with an average of 79 percent vs. 61 percent correct on all items combined. Analyses by site, gender, and ethnicity showed NHD students still outscoring peers. Both NHD males and NHD females outscored their non-NHD counterparts, and Black and Hispanic students, as well as White students, posted higher scores than peers who do not participate in NHD.

In interviews, NHD students said that reading challenging books and articles gave them the skills and confidence to tackle tough reading assignments in other subjects and synthesize large amounts of information. Students for whom English is a second language, and who are the first in their families to attend college, welcome the challenge of college reading assignments.

- **NHD students are better writers, who write with a purpose and real voice, and marshal solid evidence to support their point of view.** Even if they are not competing in essay contests, NHD students spend considerable time writing and revising—honing theses statements and exhibit information, crafting scripts for presentations and documentaries, composing process papers, and rehearsing for interviews with experts and sessions with judges.

¹² Students from all 4 sites (N=458) took the same performance assessments; these data were aggregated across sites and broken down by site, gender, ethnicity, and years of NHD participation.

Independent writing assessments show it pays off: NHD students outscored comparison-group students on both pre- and post-writing assessments, receiving more high scores (5's or 6's) on a 6-point scale, and fewer low scores. NHD essays had more sentence variety, richer vocabulary, a more authentic voice, and a more coherent and sustained argument.

Breakdowns by gender, ethnicity, state, and grade level still showed NHD students outperforming their peers.

Academic Performance

(Data sources: student scores on state standardized tests and grades in social studies, language arts, math, and science.¹³ N=1,500 for NHD students and matched comparison groups)

- **NHD participation positively affects students' academic performance in social studies.** Student scores from different assessments in different states showed that NHD students often out-perform their peers. Four years of data from the *Texas Assessment of Knowledge and Skills (TAKS)* showed that NHD students scored consistently and significantly higher than non-NHD students on the social studies assessment. In 2008-2009, for example, twice as many NHD students as achieved commended performance (87 percent vs. 37 percent). NHD students' year-end grades in social studies were also consistently higher than comparison group students'. Differences were statistically significant ($p < .001$).

On the 2008-09 South Carolina *U.S. History and the Constitution* end-of-course test, the NHD high school led the district with a 52 percent passing rate—26 percentage points above the other (non-NHD) high school in the district, 14 points above the district rate, and 9 points above the state rate. At the South Carolina middle school where NHD was required in eighth grade, eighth graders scored higher than students in the non-NHD middle school on the social studies segments of the 2008-2009 state *PASS* test, at statistically significant levels ($p < .05$). Higher percentages of the NHD students (36 percent vs. 23 percent) received an Exemplary rating.

In Colorado, high school students participating in NHD and *International Baccalaureate (IB) History* had one of the highest test scores for all IB subjects offered at the school, with the average grade of 5.02 on a 7 scale, and scored above the worldwide IB History average of 4.1.

- **Evidence also suggests that skills students gain through NHD transfer to other academic subjects.** NHD students also often outperform their peers on standardized assessments in reading, science, and math. In 2006-07, 2007-08, 2008-09, and 2009-10, the percentage of NHD students in the Texas study site who met the minimum, had commended performance, or passed the *TAKS* the first time was higher than comparison group students in reading comprehension, writing, mathematics,

¹³ Because states have different standardized assessments, data analyses were conducted by state, by metric; data could not be aggregated across sites.

and science. Differences were statistically significant ($p < .001$ level for commended performance in all subjects except writing, where $p < .01$). In Texas, NHD students' year-end grades in language arts, math, science, as well as in social studies, were also higher than comparison group students'.

In 2008-09, the South Carolina NHD school led the district with a 61 percent passing rate in English 1. Their performance was 9 percentage points higher than the non-NHD school, and 4 points above the district's passing rate. At the NHD middle school, eighth graders scored higher than students in the non-NHD middle school on the language arts, math, and science segments of the 2008-09 PASS test.

In New Jersey, historical data showed that 9- and 10- grade NHD students were performing lower than comparison-group students in language beginning in 2006-2007, but outscoring them by 2009-2010, NHD students were out-performing them. All scores show increases by year, but from 2006 to 2010 comparison-group students had a 15-point gain; NHD students had a 36-point gain.

In the Colorado middle school, more NHD students than comparison-group students consistently received Advanced performance on CSAP in writing; 57 percent vs. 42 percent (2007), 46 percent vs. 21 percent (2008); 60 percent vs. 57 percent (2009); and 53 percent vs. 32 percent (2010).

Interest in History and Civic Engagement

(Data sources: student surveys¹⁴, N=458; focus groups and interviews, N=75)

- **NHD inspires an interest in history and a deeper understanding of why it's important.** Compared to students who don't participate, NHD students are more interested in their history classes, and find those classes more interesting than their other academic classes. NHD students' interest in parallels in history, and in issues in context, is higher than their peers, at statistically significant levels. The NHD experience and the disciplinary knowledge students gain gives students a deeper understanding of why they should study history and equips them to further develop that knowledge through real research.
- **NHD students have a more mature perspective on current events.** Although survey responses suggest that both NHD and non-NHD students are fairly engaged in current events, data showed statistically significant differences in NHD students' confidence in their knowledge of history and events *not* studied in school, in their interest in staying abreast of issues, and in their skills in using and evaluating online information. NHD students' reflections on the links between past and current events also suggest that they see events in context and have a better sense of the sweep of history. Modest differences may not confirm that NHD students are better informed, but these and other findings indicate that they are more likely to seek out media and online sources to stay informed and more confident in their ability to digest and synthesize what they find.

¹⁴ Students in all 4 sites (N=458) completed the same surveys. Data were aggregated across sites, and broken down by site, gender, ethnicity, and years of NHD participation.

Cumulative and Differential Impact

- ***Trend data suggests that NHD participation has a cumulative impact over time.*** A comparison of grades to years of participation in the Texas study site showed an upward trend, or higher levels of performance with each year of participation, as did comparisons of percentages of students receiving Commended Performance on standardized assessments to years of participation. Data also suggests that students' interest in history, their confidence in research and communication skills, and their ability to interpret historical information also increase with years of participation.
- ***NHD has a positive impact among students whose interests in academic subjects may wane in high school.*** Analyses show that among Black and Latino students, NHD students still outperform non-NHD students, posting higher performance assessment scores and levels of interests and skills. Compared to non-NHD boys, and to all girls, boys participating in NHD reported significantly higher levels of interest in history, civic engagement, and confidence in research skills, on both pre- and post-surveys.
- ***Participating in National History Day increases students' college and career readiness.*** To succeed in college students need solid research and writing skills, and NHD students' more sophisticated skills and grasp of the history beyond their years suggest they are well prepared. There may not be big differences between NHD students' ratings of interest, civic engagement, and confidence and those of their peers, but the consistently higher ratings across the board may give them that edge they need to apply to schools, gain admission, and do well. These skills, along with oral communication, collaboration, and time management skills gained through NHD transfer to the workplace as well.
- ***To a greater degree than their peers, NHD students develop dispositions needed for an informed citizenry.*** NHD students are more likely to see how studying the past informs the present and the future. That they can articulate those interests suggests not only a longer view of history but also a mature ability to reflect on their learning. Analyses show that NHD's impact on students' interest in history, confidence in research and communication skills, and sense of civic engagement hold true across grade levels, gender, and race/ethnicity—and may increase with years of participation.

DISCUSSION OF FINDINGS

I. APPLIED SKILLS

What skills do students gain from National History Day, and, compared to their peers, how successfully can they apply them?

This report begins with a discussion of research, writing, and critical thinking skills that teachers say students develop through National History Day. Selecting a topic, fitting it to a theme, finding, evaluating, and synthesizing information, recasting it into an essay, play, exhibit, documentary, or website—all require sophisticated skills. Teachers also maintain that NHD students develop skills that many students don't acquire until college. Many take students to local university libraries or introduce them to inter-library loan to make it clear that this is college-level work.

Current students and NHD alumni concur: Over two-thirds of students participating in this study said NHD had had a moderate or significant impact on their ability to come up with a plan to research a topic, do research on the internet, evaluate information, and organize a report. In surveys administered in preparation for this study, an overwhelming majority of alumni say NHD had a clear impact on research and communication skills, and many volunteered examples of how they not only arrived on campus well-equipped to do college work, but continued to use the skills acquired through NHD in their careers.

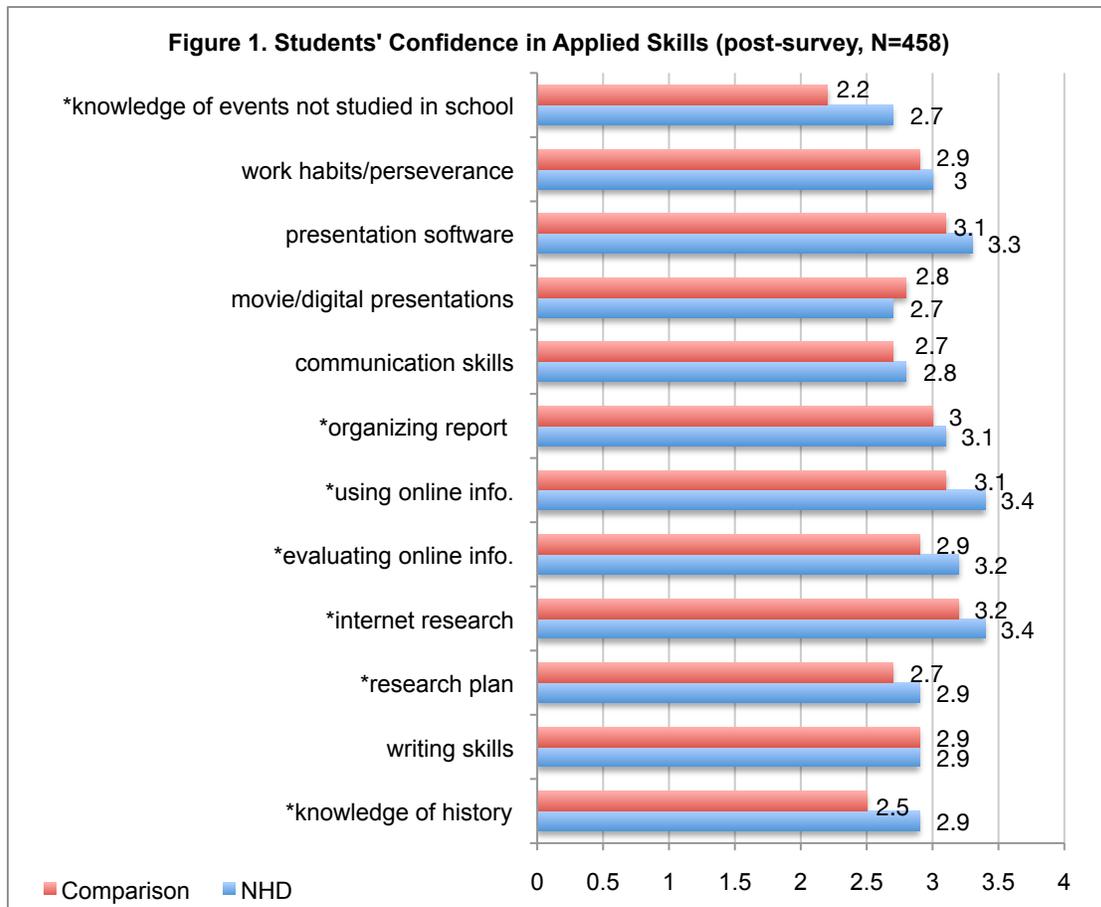
NHD students of course are not the only students who do research or long-term projects. Students write frequently, often across the curriculum. Project-based learning is a mainstay of many classrooms, popular among students and teachers. Most secondary school curriculum and standards require that students learn how to find, evaluate, and cite sources and use both primary and secondary sources. Even without the call for 21st Century skills, the sheer volume of information available online has lead teachers to spend more time teaching students about intellectual property, plagiarism, and the need to validate and acknowledge sources.

Performance assessments were designed to explore how NHD students compare to their peers. We asked both groups to rate their confidence in these skills, and apply them in a different context. ¹⁵

¹⁵ Students rated their confidence and completed performance assessments twice during the school year, or on both the pre- and post-instruments. Where the two, pre-/post- scores showed meaningful differences or trends, we have reported both. In some cases we have reported post- scores only, as a year-end performance rating for students.

Confidence in Applied Skills

Both NHD and comparison-group students reported relatively high levels of confidence in their research and presentation skills. NHD students, however, reported higher levels in 10 of the 12 skills listed. Though modest, differences between the two groups were significant at the $p < .01$ level for six skills. Both NHD and comparison-group students assigned the highest ratings to their confidence in doing internet research and using information, with means of 3.4 vs. 3.1 and 3.2 vs. 2.9, respectively (means calculated on a 4-pt. scale where 1=not at all confident and 4=very confident). The biggest differences were in students' knowledge of history and events not studied in school (2.9 vs. 2.5 and 2.7 vs. 2.2). Comparison-group students felt more confident in movie/digital presentations; ratings were the same for writing skills. Figure 1 shows the means for each group for each skill; asterisks indicate significant differences.



Source: Post-Survey

* $p < .01$

Analyses by gender and ethnicity showed NHD students still posting higher confidence levels. This pattern was repeated in other survey results, and in scores on performance assessments and standardized tests. NHD students, overall and by sub-groups, posted higher ratings or scores—not by wide margins but with marked consistency. These findings are not surprising, given that NHD and comparison-group students

were matched as closely as possible, based on school performance, but they do indicate that NHD students develop skills and confidence that may give them a competitive edge.

Application of Skills

Finding and Evaluating Information

Clearer differences between NHD students and their peers emerged when they were asked to apply research skills in closed- and open-ended performance assessment items. Almost twice as many NHD high school students correctly identified primary sources, on both pre- and post- instruments. Although high school students in both groups provided alternatives to the internet and correctly listed things that might signal a source's reliability, NHD students listed a far wider range of sources and much more detailed and sophisticated steps for establishing reliability. Most comparison-group students consistently cited the basics—books, newspapers, textbooks, encyclopedias, magazines, Wikipedia; a few also left the item blank. By contrast, NHD students listed an array of possible sources—experts, museums, lecture notes, diaries, journals, films, first-person accounts, and biographies.

Differences in students' explanations of how to judge sources were even more striking. While peers suggested checking publication date or author, NHD students said they would check for a valid copyright and reputable publisher, look for .edu—indicating a university affiliation—rather than .com sources, confirm authors' credibility by cross-checking other references, and look for corroboration across sources. NHD students also made a point of saying that they would stay away from sources like Wikipedia, which can be edited. NHD students' responses to questions about what sources should be credited also indicated that they, more so than their peers, understand that that information should not be casually appropriated.

Differences between younger or middle-school students, especially on the pre-test, were smaller: percentages of correct responses for the two items about primary sources were 94 percent and 75 percent for NHD students, compared to 61 percent and 64 percent correct for comparison-group students. The latter group gained ground over the school year, but so did NHD students, who still posted higher percentages of correct responses—100 percent and 98 percent on the two items, compared to 81 percent and 89 percent correct for non-NHD students. NHD middle-school students could more readily list sources other than the internet where they could find historical information: 89 percent correctly listed four sources, compared to 57 percent of the comparison group. When asked to describe what they would look for in determining a source's reliability, a fourth of the NHD middle-school students, compared to 10 percent of their comparison-group counterparts, correctly listed three things.

During interviews, teachers described the progressive development of research and critical thinking skills acquired through historical research. Middle-school teachers noted that, early in the process, students learn not to rely solely on the internet and Google, and are soon confidently using library databases, even from home. When seventh graders advance to eighth grade NHD projects, they know “what to look for,” and even thank seventh-grade teachers for showing them how to do annotated bibliographies. High school

teachers emphasized the importance of a research plan, “attention to detail,” and higher expectations. One teacher who demands those higher standards noted that, “If you’re a teacher of history, and you value historical research...History Day is the best thing going.”

These findings suggest that, through their research, NHD students begin to think like historians. In his research on historical thinking, and how students’ use of historical documents compares to historians’ use, Sam Wineberg found that historians use three processes: they evaluate the source, contextualize it, and look for corroboration across different sources. Students, in contrast, tend to read the documents to collect the facts. Students, says Wineberg, “read the documents as if they were historical truth; historians read them as if they were arguments.”¹⁶ Other researchers attribute students’ tendency to view the information they find as historical truth as “a lack of disciplinary knowledge and/or a lack of appropriate strategies,” stemming from the fact that students typically rely on a single textbook, a single story, and, even with multiple sources, lack strategies to “synthesize across sources” (Hynd-Shanahan, 2004, 2).

Interestingly, this research is about college students’ skills. As one NHD teacher noted, NHD doesn’t just ask students to do college-level research: it requires graduate-level skills.

Using and Interpreting Information

Both pre- and post- performance assessments asking students to interpret historical information again showed NHD students outperforming their peers. A majority of both groups answered true/false questions about historical research correctly, but percentages of correct responses were higher among NHD students, with between 10 and 20 percentage-point differences. Similarly, students drew correct conclusions or selected statements that best summarized passages, but among NHD students, percentages were higher (86 vs. 67 percent on the pre-test, and 86 vs. 72 percent on the post-test).

Middle-school students in both groups supplied correct answers to short-answer questions about a painting portraying the first Thanksgiving, but NHD students’ scores were higher: 81 percent of NHD students, compared to 64 percent of non-NHD students could identify what the artist wanted viewers to believe about the event. High school cartoons were more challenging, and some students had trouble interpreting them, but NHD students fared better. Around 40 percent of the NHD students could interpret two cartoons about women’s suffrage, correctly noting that donkeys and elephants represented political parties supporting voting rights to secure votes; just under a fifth of the comparison-group students did so. On the post-exercise, which asked students to explain cartoons about technology and social networking, more NHD students provided cogent explanations.

In interviews, NHD students said that reading challenging books and articles gave them the skills and

¹⁶ Cynthia Hynd-Shanahan, et al., “Thinking Like a Historian: College Students’ Reading of Multiple Historical Documents.” *Journal of Literacy Research* (Summer 2004), 2.

confidence to tackle tough reading assignments in other subjects: rather than being “overwhelmed by a huge textbook,” students say they know how to read through and synthesize lengthy passages. Students for whom English is a second language, and who are the first in their families to attend college, say they are up to the challenge of college reading assignments.

Overall and Grade-Level Performance

Composite scores of the various performance assessments items showed that, overall, NHD high school students were significantly ($p < .0001$) better than their peers at finding, evaluating, and interpreting historical information, with an average of 79 percent vs. 61 percent correct (see Table 5). (Because of small numbers of middle-schoolers, we used only high school students’ scores for this analysis.)

Table 5. Composite Scores on Performance Assessments

Group	Number	Percent Correct	Standard Deviation	Statistical Significance
Comparison	184	.6134	.21915	p<.0001
NHD	251	.7914	.19741	

Breakdowns of these composite scores by grade level again show NHD posting higher scores, in three out of four grade levels; scores were highest among eleventh graders. Increases by grade level suggest that gains may be due to intellectual maturation or reinforcement of skills, although the drop in eleventh grade by comparison group students does not support that view.

Table 6. High School Performance Assessment by Grade

Group	Number	Percent Correct	Standard Deviation	Statistical Significance
9 th grade				
Comparison	71	.6006	.17645	(p<.0001)
NHD	81	.7496	.16175	
10 th grade				
Comparison	43	.6312	.22584	(p<.01)
NHD	72	.7639	.23389	
11 th grade				
Comparison	30	.5952	.24909	(p<.0001)
NHD	63	.8481	.19006	
12 th grade				
Comparison	11	.8701	.13857	No stat sig.
NHD	32	.8195	.21737	

Writing Performance

In gauging students' historical thinking or their critical thinking, one good proxy is their writing skills, or their ability to construct an argument, consider or provide evidence, and maintain a narrative thread. Even if they are not competing in essay contests, NHD students spend considerable time writing and revising—honing these statements and exhibit information, crafting scripts for presentations and documentaries, composing process papers, and rehearsing for interviews with experts and sessions with judges.

Because students had a single class period in which to complete the performance assessment, we selected topics that they would likely have a ready opinion on: whether or not the legal driving age should be raised to eighteen, or whether U.S. History or World History should be taught in eighth grade (fall, pre-assessment); the pro's and con's of the impact of technology and social networking, with reference to editorial cartoons (spring, post-assessment).

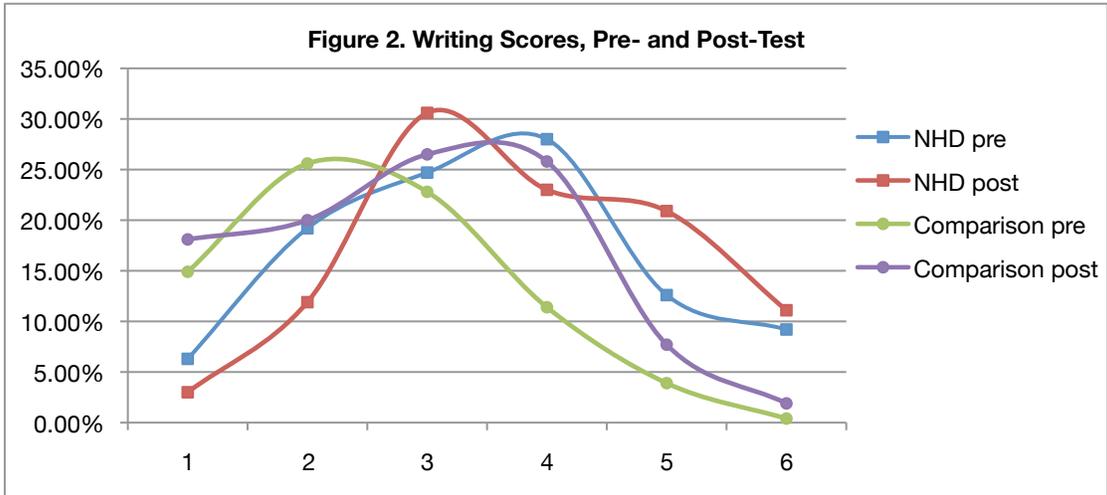
Among high school students, NHD students outscored comparison-group students on both the pre- and post- tests; both groups showed similar pre- to post- gains.¹⁷ On a 6-point scale, the mean score for NHD students on the pre-assessment was 3.5, compared to 2.5 for comparison group students. On the post-test, means were 3.8 vs. 2.9; differences were significant on both at the $p < .0001$ level. (See Appendix B, p. 53, for the rubric used for scoring.) Students' scores on both tests, and among both groups, clustered in the mid-range scores of 3 or 4, with fewer high and low scores—generally creating bell curves (see Table 7 and Figure 2).

Table 7. Percentages of Students Receiving Scores of 1-6 on Pre- and Post- Writing Assessments

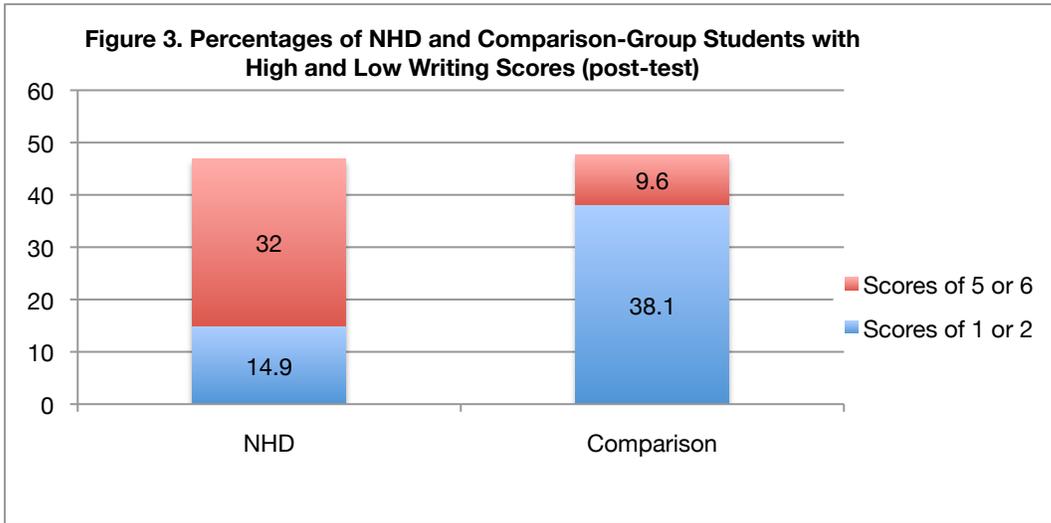
Score	NHD pre (N=252)	NHD post (N=248)	Comparison pre (N=212)	Comparison post (N=145)
1	6.3%	3.0%	14.9%	18.1%
2	19.2%	11.9%	25.6%	20.0%
3	24.7%	30.6%	22.8%	26.5%
4	28.0%	23.0%	11.4%	25.8%
5	12.6%	20.9%	3.9%	7.7%
6	9.2%	11.1%	0.4%	1.9%
Mean	3.5	2.5	3.8	2.9

Source: Pre- and post-test writing assessments

¹⁷ All student writing samples were scored by trained scorers, using a rubric created from the NAEP persuasive essay rubric (U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2007 Writing Assessment (retrieved from <http://nces.ed.gov/nationsreportcard/itmlrsx/detail.aspx?subject=writing>) and the Northwest Regional Educational Laboratory 5-Point 6+1 Trait® Writer's Rubric for 3-12 (retrieved from <http://educationnorthwest.org/resource/464>). Scorers independently scored a sample of essays at the beginning and periodically through the scoring process to establish and maintain inter-rater reliability.



What set NHD students apart from their peers were lower percentages of students receiving 1's or 2's and higher percentages of students scoring 5's or 6's: Approximately a fifth of the NHD students (22%) scored 5's or 6's on the pre-test, compared to 5% of the comparison group. On the post-assessments, almost a third, or 32 percent of the NHD students received the two highest ratings, compared to a fifth or 9.6 percent of the comparison-group students. (See Figure 3.)



Students who received higher scores were clearly writing to an audience, and a distinct voice came through. Their essays had clearer thesis sentences, a more coherent argument, and their writing flowed well from idea to idea or example to example. They generally used more sentence variety, and a richer vocabulary. Although there were strong essays in both groups, NHD students' essays exhibited these traits more often. (Interestingly, although NHD students outscored comparison-group students, their levels of confidence in their writing skills were identical, at 2.7 on a 4-point scale. NHD leaders suggested that,

because their work is held up to such scrutiny, and undergoes so much revision, NHD students may be more self-critical than their peers.)

Because writing instruction and opportunities can vary from site to site and classroom to classroom, we analyzed the data by state. (Because there may be a difference in middle and high school students' performance, we used only the latter in this comparison; Colorado was not included because we did not have a comparison group for high school students.) Results showed that the NHD students earned significantly higher scores in all three states, on both the pre- and the post-assessments.

Table 8. Writing Performance Mean Scores, by State (N=383 pre-, 328, post-)

State	Pre-		Post-		
	Number	Mean	Number	Mean	
Texas (9 th -12 th graders)	Comparison	72	2.08	56	3.05
	NHD	73	2.81****	84	3.38***
South Carolina (9 th graders)	Comparison	74	3.05	72	3.04
	NHD	74	3.70***	77	4.12***
New Jersey (9 th -12 th graders)	Comparison	54	2.20	13	1.77
	NHD	36	3.14**	26	2.65*

*p<.05; **p<.01; ***p<.001; ****p<.0001

During interviews, teachers described how NHD improves students' writing skills. Middle school teachers noted that students don't develop the more advanced writing skills in full until high school, but that NHD gives them the more basic, concrete skills needed to compose research papers: they learn how to cite sources and create a bibliography; how to look at an issue or event from a different perspective; and how to use but not plagiarize information—or how to put their ideas into their own words. According to one teacher, students often have to look up unfamiliar words, which expands their vocabulary.

Teachers at the school of the arts participating in the study described how even students who consider themselves strong creative writers reap benefits. These students, they say, are drawn to the writing options—papers, performances, scripts—and generally do well, but often have to direct their creative energies toward new types of writing. Teachers observed that they have seen more than one student who, in sixth grade, was a “good creative writer,” by tenth grade become a “good technical writer.”

Other 21st Century Skills

Although we have no comparison data or quantifiable responses, during interviews NHD teachers and students described other 21st Century skills acquired through participation.

Oral Communication and Presentation Skills

NHD students develop oral as well as written communication skills while researching and presenting their projects. First-hand accounts are often the best primary sources, and teachers encourage students to go right to the source, which could include friends and relatives, local heroes and dignitaries, government officials, and lesser-known but no less authoritative experts, such as Holocaust survivors. Teachers described particularly intimidating but rewarding interviews, such as those with cardiovascular surgeon, Michael DeBakey, and *All in the Family* producer, Norman Lear (which eventually led to a summer internship). Students also describe the skills and confidence they gain from being interviewed by judges, who not only ask hard questions, but may also deduct points, not necessarily for students who are not eloquent, but for those who are not prepared or who rely too much on fellow presenters.

Collaboration

Judges' insistence that all team members be able to answer questions is part of the collaborative model defined and required by NHD. Although students undertake tasks that match their strengths and interests, in NHD projects one or two students cannot do all the work. As noted earlier, projects are common in many schools, popular with students who enjoy working with peers. NHD teachers maintain that what sets NHD apart from most project work is that all students have to immerse themselves in the topic and contribute. One student pointed out that collaboration is not just a requirement, but that it is the "History day spirit: to help each other." Students at the school of the arts, who routinely perform, say that NHD creates "a different group dynamic." In other projects, "teachers direct," but NHD is more self-directed, and "shows their personality more."

Time Management, Problem-Solving, and Perseverance

When asked what they would say about NHD to a sibling or friend considering participation, one high school student said: "don't procrastinate." Teachers unanimously agree that time management is one of the most important and practical skills students learn. One teacher compared the process of doing NHD to "preparing for a marathon...you can't do the work overnight or at the last minute."

Students learn time management by completing the requirement of NHD—gathering primary and secondary sources, writing process papers, creating their exhibits, performances, documentaries, and web sites—but, perhaps more important, they continually refine their work, often many times over the course of the year. They incorporate peers' suggestions, new research, feedback from teachers, and—if they enter competitions—feedback from judges. This long-term effort may be what sets NHD apart from other project-based learning activities, and what helps students develop not only time management skills but also persistence. When asked how he would describe NHD to a friend or sibling, a middle-schooler preparing to revise once again accurately said, "When you think you're done, you're not."

II. ACADEMIC PERFORMANCE

Does National History Day have a positive impact on students' performance on high-stakes tests—not just in social studies, but also in other academic subjects?

Applied skills are solid indicators of college and career readiness, but school administrators still look to high-stakes test scores to confirm program effectiveness. Although test scores may be the holy grail of validation, the link between program and performance is often hard to prove: much else goes on in a school day, particularly for secondary students who take multiple classes with different teachers.

The variability in NHD treatment adds another challenge to proving causation. We initially thought we could control for treatment by creating an index for teachers' implementation. That proved difficult because NHD is, in some cases, part of the curriculum and classroom activities; in others, it is still mandatory, but work is done outside of class. NHD can also be an extra-curricular option. There were also challenges related to students' exposure. Some students start their NHD projects at the start of the school year, some wait until second semester, and some start the previous summer and work right up until contests. Our sample sizes were often too small to separate students by grade level and years of experience; moreover, we had eighth graders with two years' experience, and eleventh or twelfth graders with one year's experience.

We therefore cast a wide net to explore the program's impact on students' academic success, requesting, for three years (2007-2008 through 2009-2010) from all four sites, scores on standardized tests in all academic subjects,¹⁸ grades or GPAs, and other performance data such as Advanced Placement (AP) or International Baccalaureate (IB) scores. We also requested attendance, behavioral referral, and suspension data, and students' gender, ethnicity, and special status, such free- or reduced-lunch, Special Education, or English Language Learner. In some sites, sites were not allowed to identify students by special status, and in some cases not all data was stored in the current database. Student mobility posed an additional challenge, since district databases did not include all records of transfer students.¹⁹

The findings in this section are reported on a state-by-state and test-by-test basis. Aggregating data was not possible because states administer different standardized tests; even within a state, assessments may

¹⁸ These included the Texas Assessment of Knowledge and Skills (TAKS), the New Jersey Assessment of Skills and Knowledge (NJASK), the South Carolina's Palmetto Assessment of State Standards (PASS) and Palmetto Achievement Challenge Test (PACT), and the Colorado Student Assessment Program (CSAP).

¹⁹ We also requested data on attendance, behavior referrals and suspensions, student mobility rates, and graduation rates—all of which NHD could affect, and all of which that might be a factor in or reflect students' academic performance. Not all sites were able to provide the data, because it was not housed in the same database as other performance data or otherwise was not easily retrievable. High levels of student mobility also meant that there was considerable missing data. As a result, we often did not have enough data to conduct meaningful analysis. In the few cases where we did have sufficient data, we did not find any clear differences between NHD and comparison-group students.

change from year to year. Students—especially high school students—are seldom tested annually, on the same subjects. There are more findings from Texas because Texas administers multiple tests in multiple subject areas, on an annual basis. The state puts a strong emphasis on data, and thus has a robust data system, readily accessible by districts. NHD also has a long history in Texas, and our sample there was often larger and spanned more grade levels.

Social Studies Performance

Social studies is the content area where we would most likely find evidence of impact. The findings below include students' performance on annual tests, end-of-course exams, AP and IB exams, and grades. These bullets, and those for other academic areas, do not report all the assessments and scores we compared. There were cases where we found no significant differences between NHD and comparison-group students, or year-to-year or subject-by-subject comparisons showed no clear differences or trends.

Texas

- Four years of data, from 2006-10, showed NHD students out-performing matched comparison groups of peers on the Texas Assessment of Knowledge and Skills (TAKS) social studies test. In 2008-2009, 87 percent of the NHD students achieved commended performance, compared to 37 percent of the comparison-group students; in 2009-2010, 73 percent of the NHD students received the highest rating, vs. 53 percent of the comparison-group students. Higher percentages of NHD students also met the minimum proficiency level or passed the test for the first time. Differences were statistically significant ($p < .001$).
- NHD students' year-end grades in social studies were higher than comparison group students'. Differences were statistically significant ($p < .001$).

South Carolina

- On the new 2008-2009 U.S. History and the Constitution end-of-course test, the NHD school led the district with a 51.6 percent passing rate. Students' performance was 26.1 percentage points above the comparison school's scores, 14 points above the district passage rate, and 9.2 points above the state rate.²⁰ The test counts for 20 percent of students' final grades.
- At the South Carolina middle school where NHD was required in eighth grade, eighth graders had Social Studies scale scores on the 2008-2009 PASS test that were significantly higher ($p < .05$) than the comparison group's scores, with a mean of 640.09 versus 623.19. Higher percentages of the NHD students (36 percent vs. 23 percent) received an Exemplary rating.
- In 2010, when students did not participate, percentages of students receiving an Exemplary rating

²⁰ In addition to school data and sites interviews, findings come from an article by Robert Monnie, "End-of-Course Scores Released," *The Link*.

went down, from 36% to 30%, and percentages of students who did not meet the proficiency level went up steeply from 18% to 40%. Table 9 show the mean 2009 and 2010 PASS scores for both groups of students; Table 10 shows the percentages of students who did or did not meet the proficiency level, and the percentages receiving an Exemplary rating.

Table 9. NHD and Comparison Groups' Mean Scores, PASS test in Social Studies, 8- Grade

2009	Number	Mean (SD)
Comparison	74	623.19 (45.9)
NHD (participation mandatory)	94	640.09 (47.5)
2010	Number	Mean (SD)
Comparison	73	625.97 (51.9)
NHD (no participation)	89	623.52 (52.7)

Table 10. NHD and Comparison Groups' Performance Level, PASS test in Social Studies, 8- Grade

2009	Comparison	NHD (participation mandatory)
Not Met	31.1% (23)	18.1% (17)
Met	45.9% (34)	45.7% (43)
Exemplary	23.0% (17)	36.2% (34)
2010	Comparison	NHD (no participation)
Not Met	34.2% (25)	40.4% (36)
Met	35.6% (26)	29.2% (26)
Exemplary	30.1% (22)	30.3% (27)

Source: Chesterfield County South Carolina school district

Colorado

- In Colorado, high school students participating in NHD and *International Baccalaureate (IB) History* posted some of the highest test scores compared to all IB subjects offered at the school, with the average grade of 5.02 on a 7-point scale, and scored above the worldwide IB History average of 4.73.
- Social studies grades were also higher among NHD students, but the small sample size limited tests for significance.

Performance in Other Academic Areas

Analyses of students' performance outside of social studies also suggested that the skills students gain through NHD can transfer to other academic subjects.

Texas

- From 2006-2010, the percentage of NHD students who met the minimum, had commended performance, or passed the Texas Assessment of Knowledge and Skills (TAKS) the first time was higher than comparison group students in reading comprehension, writing, mathematics, and science. Differences were statistically significant ($p < .001$ level for commended performance in all subjects except writing ($p < .01$). In reading, science, and math, close to two-thirds of the NHD students received commended performance, compared to between 12 and 25 percent of comparison-group students.

- NHD students' year-end grades in language arts, math, science, as well as in social studies, were higher than comparison group students'. Differences were statistically significant ($p < .001$ for math and science, and $p < .01$ for language arts).

South Carolina

- On the English end-of-course test, added to the state's accountability system in 2005-2006, the NHD school led the Chesterfield County School district with a 60.6 percent passing rate in English 1. The school's performance was 8.8 percent higher than the comparison school, and 4.1 above the district passing rate. The test counts for 20 percent of students' final grades.²¹
- Cheraw, the NHD school, was in the top ten of all public high schools in the state on passage of AP exams in 2008-2009: 29 students took 33 AP exams; 28 percent received 3's or higher. In 2008-2009, 18 comparison-group students took 29 exams, with a 21 percent passing rate.²²
- There were no significant differences in the High School Assessment of Proficiency (HSAP) Language Arts performance in 2008 and 2009 between tenth graders (the year the annual gateway test is administered the first time) in the two schools. In 2009, comparison-group students scored significantly higher on the math portion of the assessment.
- At the NHD middle school (where participation was mandatory until 2009-2010) eighth graders scored higher than students in the non-NHD middle school on the Language Arts, Math, and Science segments of the 2008-09 PASS test. On the 2009-2010 test, percentages of students in the (former) NHD school who received exemplary ratings in Language Arts went down slightly, and the percentage of students who did not meet the proficiency requirement went up.

New Jersey

- New Jersey students begin NHD participation in ninth grade. A look at historical data showed that 9- and 10- grade NHD students were performing lower than comparison-group students in Language beginning in 2006-2007, but scores rose. In 2009-2010, NHD students were out-performing the matched comparison group. All scores show increases by year, but from 2006 to 2010 comparison-group students had a 15-point gain; NHD students had a 36-point gain. (See Table 11.)

Table 11. NJ ASK Language Scores prior to and after NHD Participation

Year	Comparison (n=26)	NHD (n=18)
2006-07	201.88	192.07
2007-08	207.35	199.80
2008-09	211.72	201.88
2009-10	216.71	228.00

Source: Paterson School District

²¹ In addition to school data and teacher interview, sources include Robert Monnie, "End-of-Course Scores Released" (*The Link*, no date available), p. 1.

²² School data, teacher interview, and Robert Monnie, "Cheraw Excels in AP Exam Results" (*The Link*, Dec. 22, 2009), 1 and 8A.

Colorado

- In the Colorado middle school, more NHD students received Advanced performance on CSAP in writing; 57 percent vs. 42 percent (2007), 46 percent vs. 21 percent (2008); 60 percent vs. 57 percent (2009); and 53 percent vs. 32 percent (2010). (Reading scores showed no consistent differences between NHD and non-NHD students.)
- Although GPAs in the CO middle school was relatively high, NHD students' 2010 grades were higher, and differences were significant in English, math, and science. (ELA: M=3.7 vs. 3.2, $p<.01$); Math: M=3.6 vs. 3.25, $p<.05$; Science: M=3.28 vs. 2.79, $p<.05$.)

III. INTEREST & ENGAGEMENT

How does NHD students' interest in history, and their perspectives on current events, compare to their peers'?

A wealth of anecdotal evidence from teachers, students, and parents, indicates that NHD students develop a genuine interest in history. Some students start with an interest in history, but not all students have that passion. During interviews, NHD teachers explained how they engage students “not necessarily fascinated with the human story over time” by suggesting they start with something they do care about: At a South Carolina teacher’s urging, girls interested in style researched Madame C.J. Walker, and created an award-winning project entitled “‘Hair Works’: An Innovation that Put Black Women to Work.”

We used surveys not only to gauge students’ interest in past and current events, and their ability to see the links between the two, or that narrative thread alluded to by the Colorado student.

Interest in History

Students’ responses to two post-survey items show that NHD students in all four sites are interested in their current history classes—and often find them more interesting than their other academic classes. Comparison group students were also fairly interested, just not as interested as their NHD counterparts. Differences were statistically significant in 3 of the 4 sites. Figures 4 and 5 below show the relative means for both groups, on a 4-point scale. (For the first item, 1=never interested and 4=almost always interested; for the second, 1=a lot less interesting, and 4=a lot more interesting.)

Figure 4. Students' Interest in History Classes

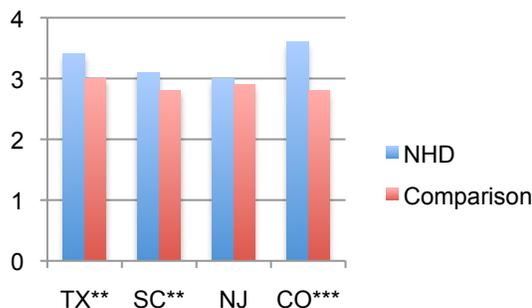
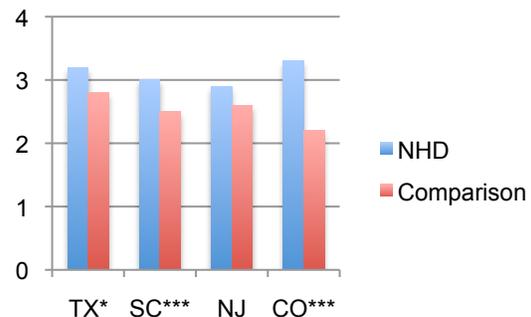


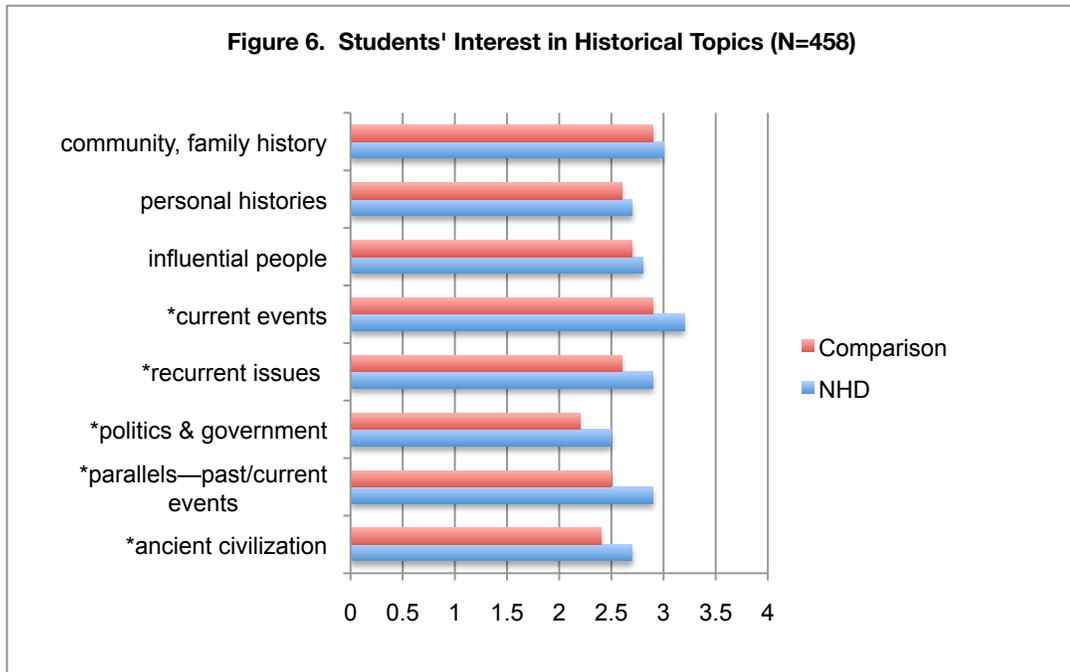
Figure 5. Students' Interest in History Classes Compared to Other Classes



*p<.05; **p<.01; ***p<.001
Source: Post-Survey

Comparing his interest in history, or NHD, to his interest in history class, one Colorado student explained his preference for the former because in classes, “what gets left out is the stories.”

Both pre- and post-surveys also asked students about their interest in historical periods and topics. (Figure 6 below shows post-survey means, on a 4-point scale where 1 = not at all interested; 4 = very interested.)



*p<.001
Source: Student Post-Survey

Again, all students, NHD and non-NHD alike, expressed interest in various topics, but interest was stronger among the NHD group, especially on the post-survey (not necessarily because of pre-to post increases, but because interest among the comparison group declined more from pre- to post). Differences were statistically significant for five of the eight topics.

More differences emerged in students' appreciation of history, as reflected in their comments on what historical topics interested them. After rating their interest in various topics, students were asked to look back over their ratings to see whether they were consistently low or high, or mixed, then to elaborate on their interests, based on their ratings and any patterns they observed. Researchers coded students' responses in two ways: first, for whether they reflected a low, high, or mixed interest; and, second, for whether they included details to explain their ratings, and genuinely reflected on their interests, or whether they supplied only short answers with few details. The analysis showed that:

- Most NHD students (70 percent) expressed high interest, and only 10 percent suggested little interest. In contrast, 34 percent of the comparison-group students expressed high interest, and a number of students left the item blank.

- There were more comparison group students who noted a single period (e.g., the Holocaust) or interest (family history).
- NHD students were more apt to explain their ratings, by margins of 3 to 1. In explaining their views, comparison group students tended to say, “Some things are more interesting than others in history, or “Some topics just pull you in more than others.”

It did not come as a surprise that the students involved in a year-long history project, many of whom had explained their research to judges at competitions, would be more likely to share their interests. What was notable was how reflective the NHD students were, and how often their explanations included references to the links between past and present. The previous section cited studies on historical thinking, or the dispositions that students who study history develop or that historical study requires—the “specialized ways of knowing and thinking, habits of mind and cognitive processes that typify historians’ approaches to the past.”²³ These habits of mind evolve from chronological thinking to considerations of context to the event’s connection to larger trends. NHD students’ comments suggest that they are fairly evolved historical thinkers:

I believe that knowing where you came from and where you're going is essential to succeed in society. By knowing what has happened in the past, and analyze what is presently occurring we can change what will happen.

Big picture history is more interesting for me, especially when looking at parallels between current events and past events...names and dates don't mean much to me, but ideas and patterns have a huge influence.

I'm interested in History because I love learning the struggles of the land and what people went through to get where we are today.

I love seeing how history repeats itself and looking at how governments and daily events have changed, so I'm interested in parallels in history and in current events as well as how the government works. The rest of them spark a little interest, but only a little just because I'm more intrigued by the way things are affected as a whole rather than little by little bit.

I am interested in history because past events have influenced present day life and it's important to know the exact events that made something the way it is now. History is very broad and it's fun to learn about someone else's history and compare it to yours or your people's.

²³ From the *National Standards for History*. rev. ed. (Los Angeles: National Center for History in the Schools, 1996).

Civic Engagement

Understanding the links between past and current events can inspire a sense of engagement in local and national issues, as can the projects that take NHD students out into their communities—to explore a university library, museum, or other archived collections, or to talk with community members with first-hand knowledge of a topic. Students gain a respect for local history, and the part it played in a larger issue or arena, and a better sense of the issues, events, and stories that define a community.

To look at this important aspect of NHD participation, and see how NHD students' perceptions compared to their peers, we included survey items related to civic responsibility and engagement.²⁴ Responses to survey questions about civic engagement did not show major differences between NHD and comparison-group students. Most students assigned relatively high ratings to various items—as a whole these students are fairly interested, engaged, and confident. NHD students' ratings were, however, consistently higher. Composite means for the pre- and post-survey show a statistical difference at the $p < .01$ level (see Table 12; means calculated on a 4-pt. scale where 1=strongly disagree and 4=strongly agree.) This is the pattern repeated in students' responses to survey questions about their interest in history (and confidence in skills).

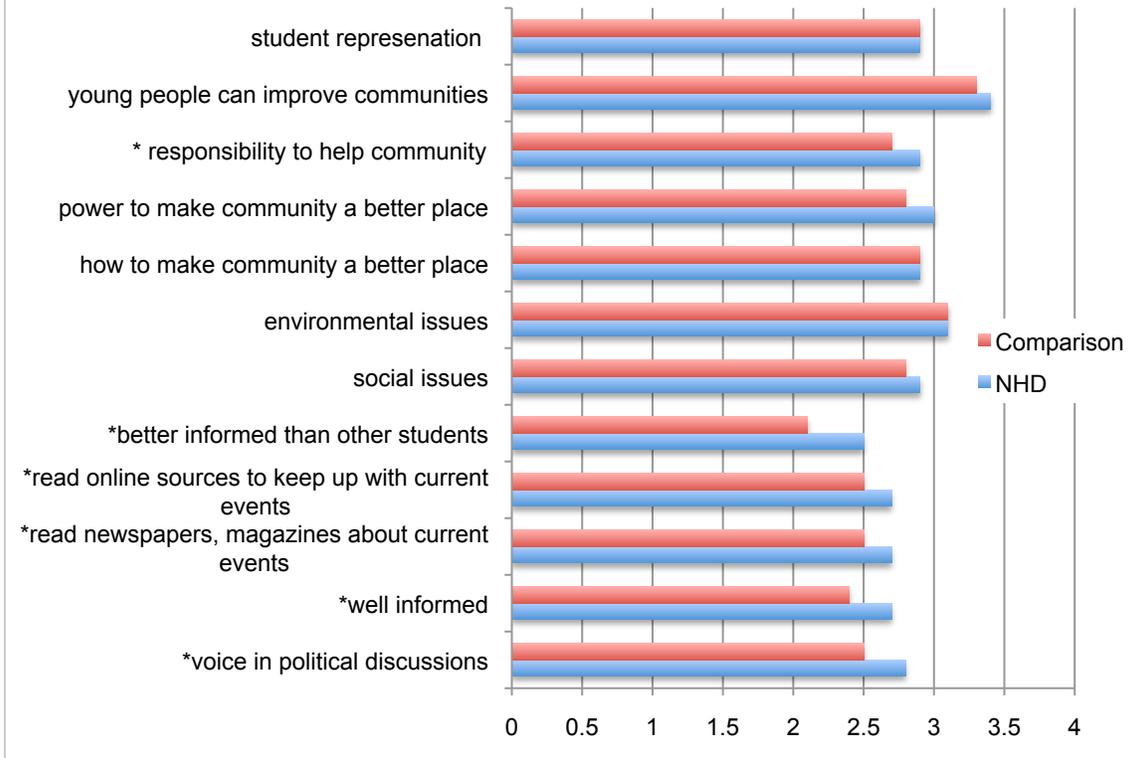
Table 12. Composite Means for Pre- and Post-Survey Items on Civic Engagement

Group	Number	Mean	Significance
Comparison Group, pre	158	2.62	p<.01
NHD, pre	249	2.86	
Comparison Group, post	156	2.68	p<.01
NHD, post	249	2.90	

Figure 7 shows post-survey means for the individual items. Both NHD and comparison-group students most strongly agree that young people can improve communities ($M=3.4$ and 3.3 respectively), and care about environmental issues ($M=3.1$ for both groups). Responses to items related to students' self-assessments of how informed they are, and how they stay informed, showed the biggest differences between the two groups. (Asterisks indicate statistical significance at $p < .05$ or less.)

²⁴ Items adapted from Tufts National Survey of Civic Participation, <http://activecitizen.tufts.edu/downloads/questionnaire.pdf>; and UC-Berkeley, Service-Learning Research and Development Center, survey of civic responsibility, http://www.peecworks.org/PEEC/PEEC_Inst/0038CEB0-007EA7AB.1/furcosurvey%20civicrespons.pdf.

Figure 7. Students' Engagement in Current & Civic Issues & Events (N=458)



Source: Student post-survey

While the data does not show that NHD students are better informed or more engaged in current events or issues than their peers, it does suggest that NHD students are more likely to do the things that keep them informed and involved. During interviews, some students confessed that their projects left little time for keeping up with current events, but comments also suggested that, when they do, their NHD skills and experiences serve them well: as one student noted, “When I watch the news, I connect the dots,” further explaining that he links current reports to what he already knows about a country or how an issue played out in a different context.

IV. DIFFERENTIAL & CUMULATIVE IMPACT

Does NHD participation have a positive impact on all students, and does impact build over time?

NHD's Impact on Different Kinds of Students

During interviews, NHD teachers repeatedly said that all types of students participate—and excel—in the program. A Colorado teacher said that NHD was “the finest standards-based curriculum for all students,” stressing the “all.” Teachers say that NHD is also an “excellent” way to keep students motivated in school, because they can pursue what interests them.

Teachers also shared examples of special education students who gained confidence and skills with demanding projects, shy students who emerged from their shells, disinterested students who found new motivation in researching topics they chose, students from war-torn countries who found the challenges of NHD mild in comparison, non-native speakers whose participation eased the transition into an American school—for students and families alike. Teachers and students described the inclusiveness and mutual support that are part of NHD. Teachers also described the social benefits of group projects, for students who may have been “outcasts” but make new friends and show group partners a “different side” of themselves; and for students who may have been “intimidated by the brilliance of their peers” but become the “resident expert” on a particular aspect of a NHD topic. As one teacher explained,

We have students of all ability levels participating. I have a student who is classified as a 504 student—she is legally blind—working on a documentary. While we have had to adapt many things for her, she has a wonderful feel for what should be included. I've had many students who feel their accomplishments in NHD, no matter what level they progress to, are the pinnacle of their high school careers. I've had students who were failing classes rise to the occasion because they must be passing to compete. Some students change their entire view of school as a result of participation in the program. One young man went from failing regular classes to excelling in AP classes as a result.

Collecting hard data on the academic and social benefits of NHD for different kinds of students was beyond the scope of the study. We did, however, analyze data by various background factors, to the extent that school districts were able to provide demographic data, and samples were large enough to separate students into different cells based on gender, ethnicity, and special designation.

Applied Skills, by Gender and Ethnicity

- Composite scores on the performance assessments show that both NHD girls and NHD boys out-score peer groups, at statistically significant levels.
- Among Black and Hispanic NHD students, as well as White students, also out-score peer groups, at statistically significant levels.

Table 13 shows the breakdowns by gender and ethnicity. (Composite scores were calculated for high school students only, because of the low sample size among middle-school students.)

Table 13. High School Performance Assessment By Gender and Ethnicity

	Number	Percent Correct	Standard Deviation	Statistical Significance
<i>Male</i>				
Comparison	81	.6305	.21032	(p<.0001)
NHD	133	.7863	.21646	
<i>Female</i>				
Comparison	68	.6292	.22697	(p<.0001)
NHD	108	.7837	.18745	
<i>White</i>				
Comparison	44	.6331	.20120	(p<.0001)
NHD	89	.8443	.15212	
<i>Black/Latino</i>				
Comparison	86	.6487	.22593	(p<.01)
NHD	133	.7428	.22097	

Source: Student performance assessment

- Breakdowns in the writing performance scores by gender and ethnicity again showed both male and female students posting higher scores than their peer groups, on both the pre- and post-assessments, at statistically significant levels. (See Table 14.)

Table 14. Writing Performance by Gender and Ethnicity

Ethnicity		Pre		Post	
		Number	Average Score	Number	Average Score
<i>White</i>	Comparison	44	3.15	41	3.22
	NHD	88	4.10*	88	4.34*
<i>Black/Latino</i>	Comparison	77	2.18	71	2.87
	NHD	118	3.10*	119	3.56*
Gender					
<i>Male</i>	Comparison	118	2.62	75	3.08
	NHD	137	3.61*	123	4.00*
<i>Female</i>	Comparison	85	2.34	54	2.70
	NHD	106	3.30**	101	3.57**

Source: Student writing assessment

* p<.001; ** p<.0001

Academic Performance

- Our analysis of differences based on gender or ethnicity showed that, where, overall, NHD students out-performed comparison-group students, findings were consistent for sub-groups: male NHD students outperformed non-NHD males, and female NHD students outperformed non-NHD females. Similarly, Black and Hispanic NHD students outperformed their peer groups, as did white students.
- The differences in student outcomes were also very similar, regardless of gifted/talented status: across sites, our sample of students designated as gifted students was small, and gifted students, NHD or comparison, were performing at relatively high levels. Where comparisons showed NHD students outperforming non-NHD students, this was true of non-gifted sub-groups as well.
- Extremely small samples meant that we could not compare academic performance based on Special Education or Limited English Proficient status.

Interest and Confidence

- Composite scores show that, compared to their peer groups, NHD students have higher levels of interest and confidence, regardless of ethnicity.
- An examination of interest, engagement, and confidence data by gender, showed that, compared to comparison-group boys, and to all girls, boys participating in NHD reported significantly higher levels, on pre- and post-surveys. Interestingly, levels declined from pre- to post-, which may be due to end-of-school doldrums. (For NHD students, it may have something to do with the middle-schooler’s comment about NHD and preparing for a contest: “When you think you’re done, you’re not.”) Girls

taking part in NHD also reported higher levels of interest, engagement, and confidence than non-NHD girls, but differences were less dramatic. Table 15 shows the means and significance levels.

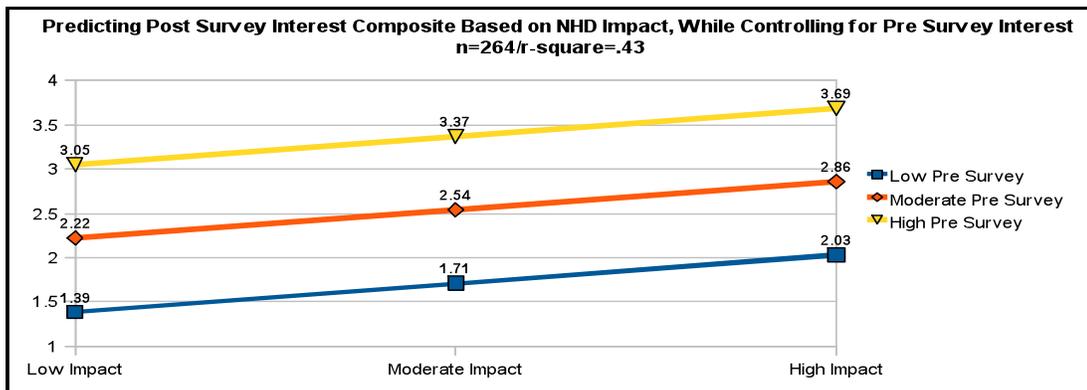
Table 15. Differences between NHD and Comparison Group Students Based on Gender and Levels of Interest, Engagement, and Confidence

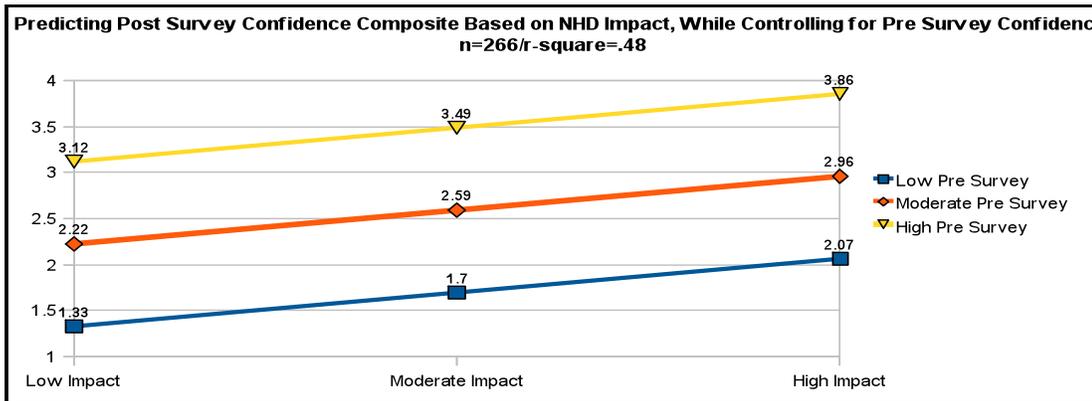
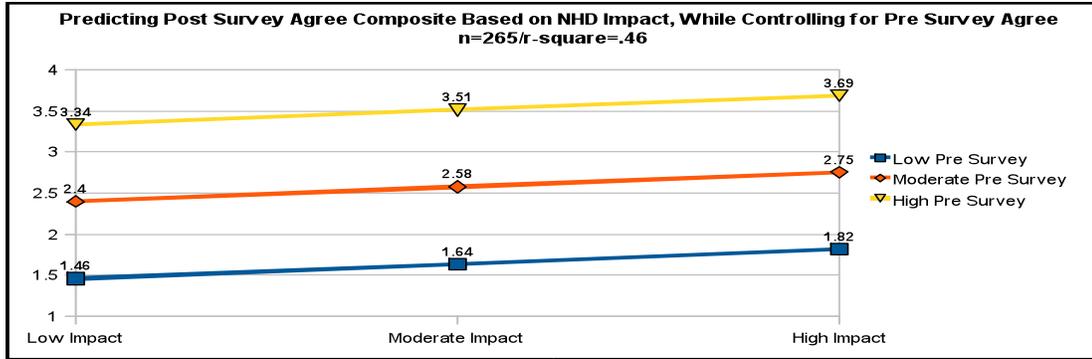
	MALE		FEMALE		
	NHD	Comp	NHD	Comp	
Interest Composite	Pre-Survey	3*	2.85	2.78	2.81
	Post-Survey	2.92***	2.64	2.76	2.61
Engagement Composite	Pre-Survey	2.99***	2.78	2.83	2.7
	Post-Survey	2.94**	2.77	2.84*	2.64
Confidence Composite	Pre-Survey	3.11**	2.97	3.01*	2.95
	Post-Survey	3.09*	2.88	2.97*	2.78

Source: Student Surveys
 *p<.05; **p<.01; ***p<.001

Data also suggest a positive link between students' interest in history, civic engagement, and confidence—and their perceptions of impact. Students reporting that NHD has had a greater impact also reported more interest in past and current events, and more confidence in their research skills. Figure 8 shows the regressions graphs for all NHD students.

Figure 8. Correlations between Impact and Interest, Engagement, and Confidence



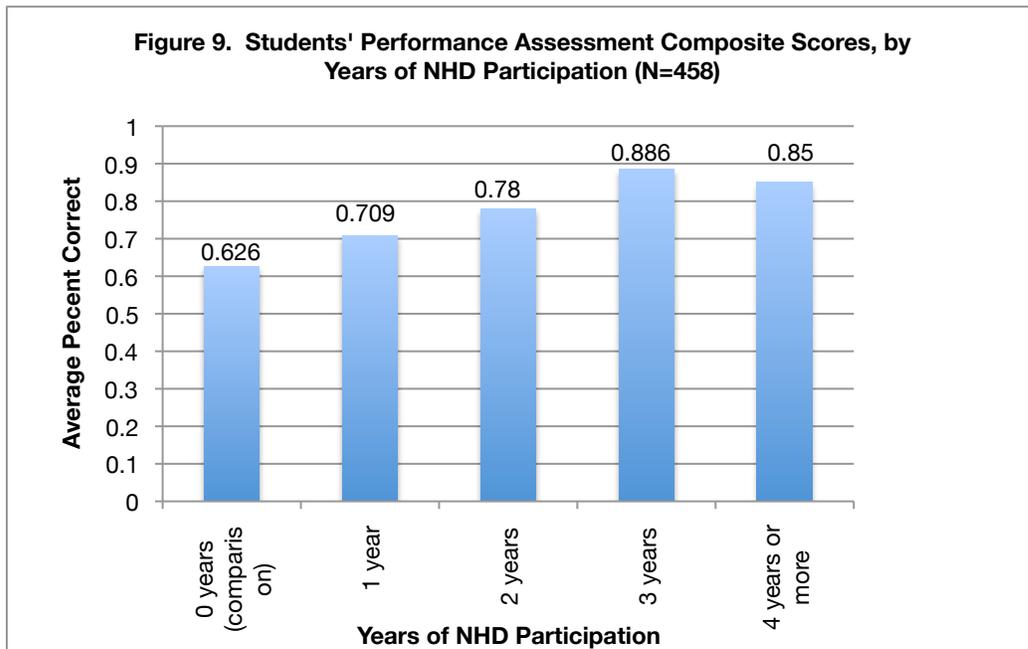


Impact over Time

Students can participate in NHD for multiple years, from sixth through twelfth grades, and many do. To see whether participation might have a cumulative impact on students, we analyzed data from each component of the study—applied skills, academic performance, and interest and engagement—based on years of participation. It should be noted that the years of participation do not always correspond to grade levels, so upward trends do not necessarily reflect intellectual development. A student's first of participation could be in sixth grade or twelfth grade, or anywhere in between; students with single or multiple years of participation could be in eighth through twelfth grades (no sixth or seventh graders took part in the study).

Applied Skills

- Composite scores on the performance assessments testing students' ability to identify and evaluate sources and interpret historical information increased steadily with each year of NHD participation, up to three years; the average percent correct dropped slightly after four or more years (see Figure 9).



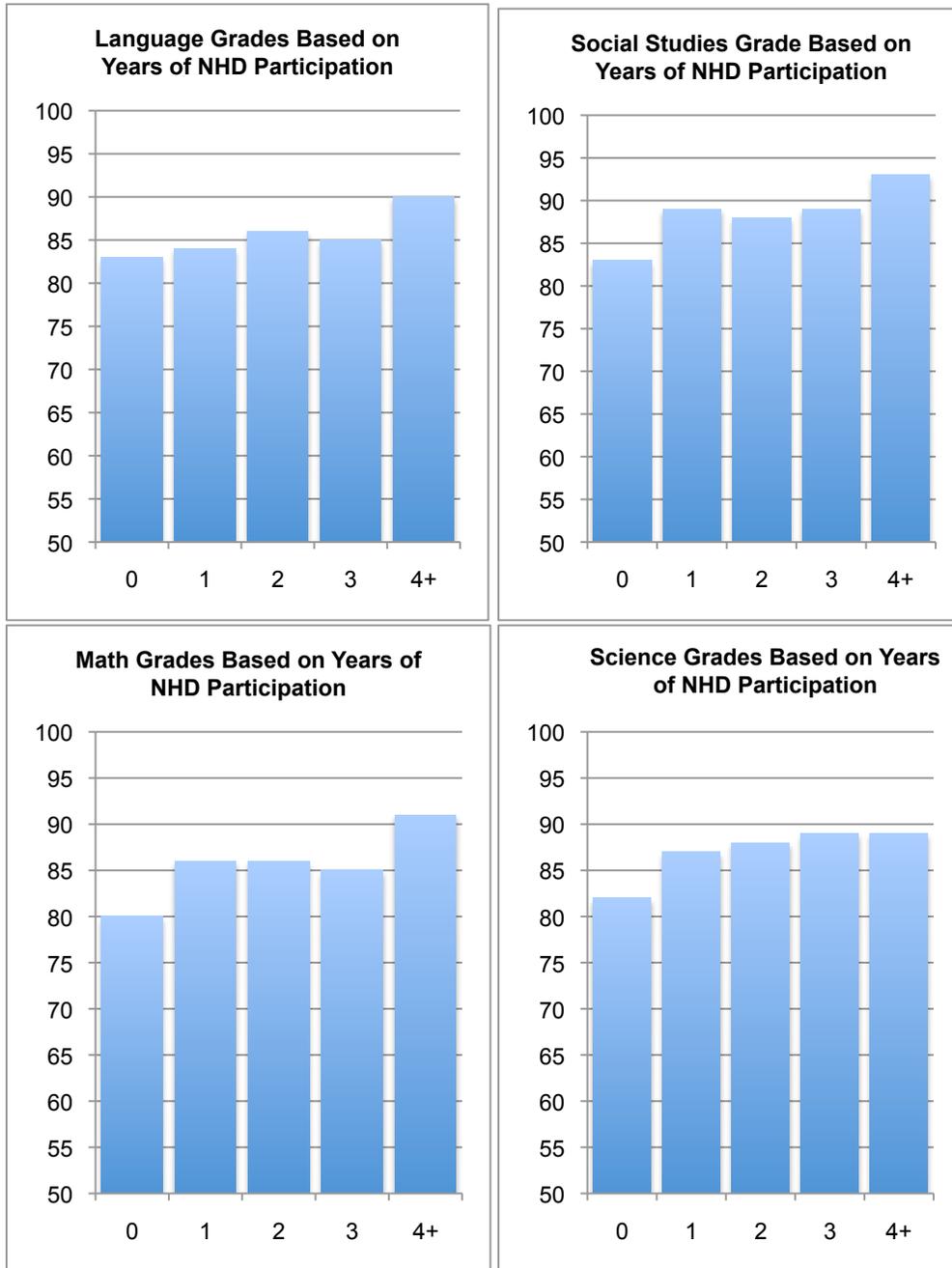
Source: Student Performance Assessment, Composite Scores

Academic Performance

Years of NHD participation also appeared to be linked to upward trends in school performance.

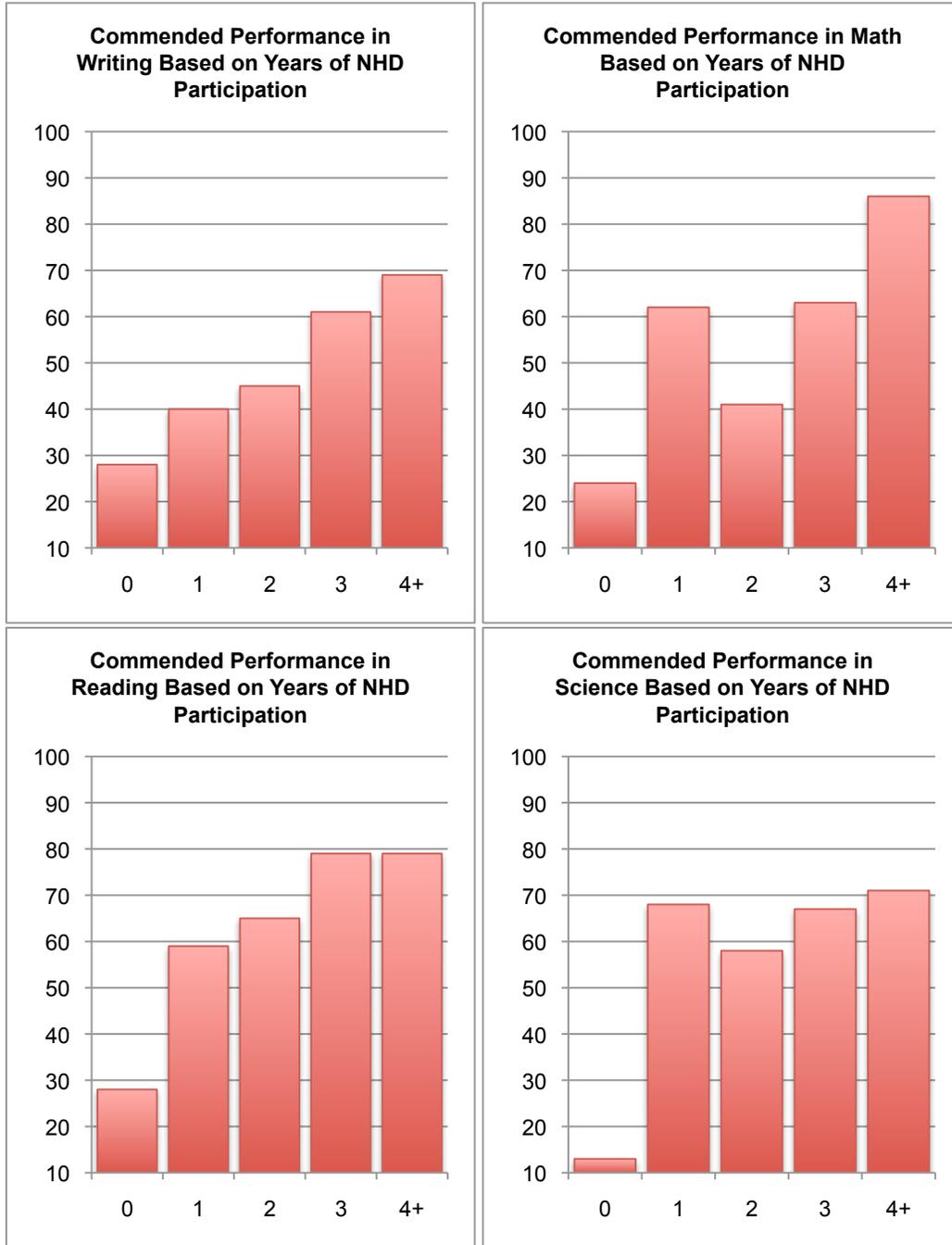
- A comparison of grades to years of participation showed an upward trend, or slightly higher levels of performance with each year of participation. Figure 10 shows Texas students' 2009 grades in language, social studies, math, and science, based on years of NHD participation; 0 years represents the comparison group. (Comparisons of academic performance and years of participation include Aldine, Texas students only because only Aldine stored sufficient data, over time, to conduct analysis.)
- A comparison of percentages of students receiving Commended performance in writing, math, reading, science, and social studies to years of participation also suggests upward trends or a spike after one year of participation. Figure 11 shows the trends by subject area; Figure 12 combines writing, reading, and social studies in a single graph; 0 years represents the comparison group.

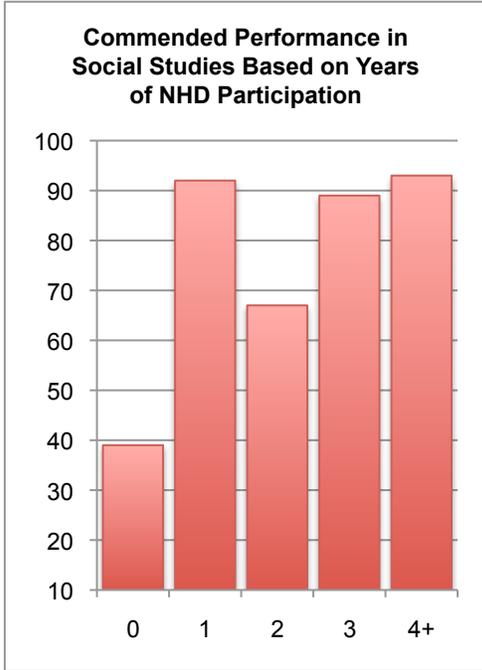
Figure 10. Grades Based on Years of NHD Participation



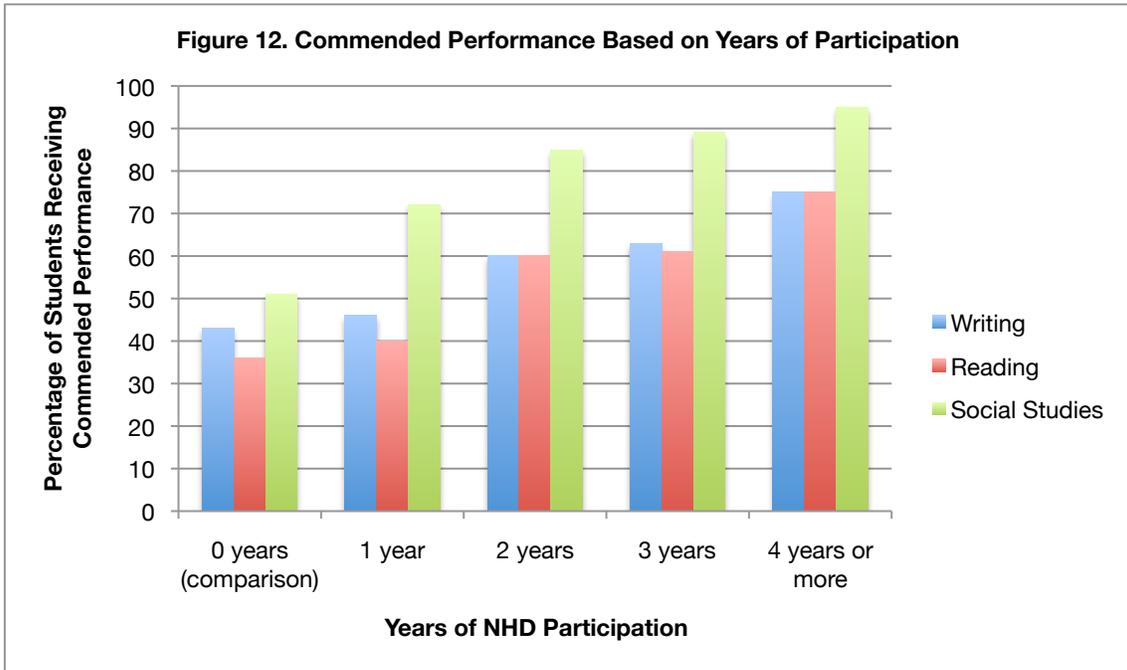
Source: Aldine Independent School District, 2009 school data

Figure 11. Percentages of Commended Performance on TAKS Based on Years of NHD Participation





Source: Aldine Independent School District, 2008 TASKS data

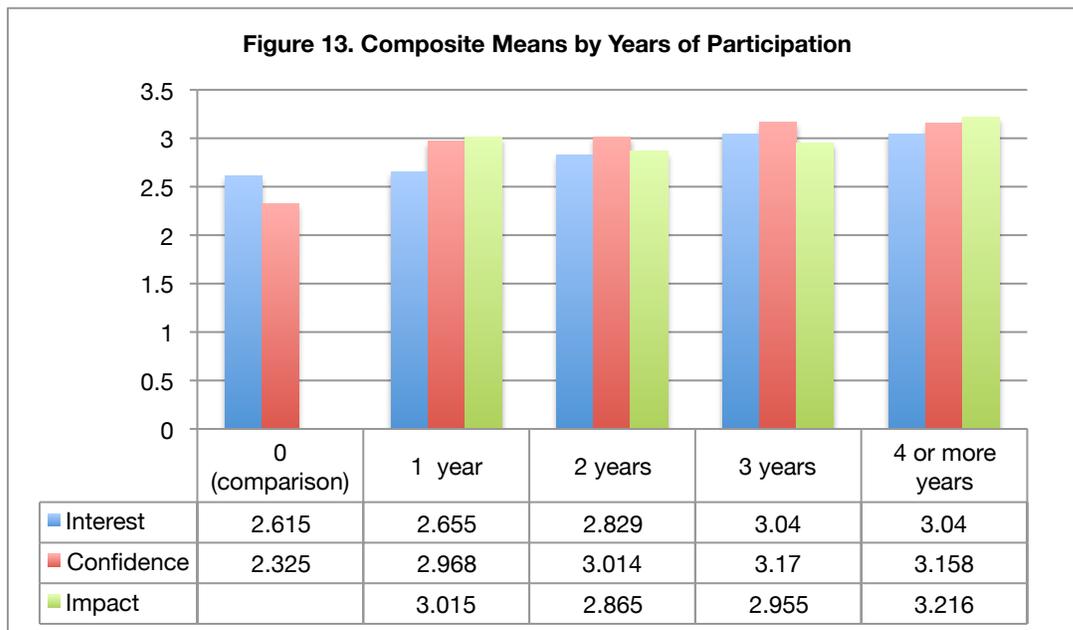


Source: Aldine Independent School District, 2008 TASKS data

Interest in History and Civic Engagement, and Perceptions of Impact

- For the interest composite, based on students' ratings of their interest in various historical periods and topics, there was a slight positive change with each year of NHD participation. Comparison group students were significantly lower ($p < .05$) than students with 2 or more years participation in NHD.
- Composites for students' confidence showed that NHD students were more confident in their skills than their comparison-group counterparts, and that confidence builds slightly from year to year. Differences were not statistically significant.
- When we formed composites for NHD students' ratings of the impact of the program on research and communication skills, we found a statistically significant difference from 1 year to 4 years or more (comparison-group students were not asked about impact). These differences were statistically significant at the .01 level.

Figure 13 shows the changes in levels of interest, confidence, and perceived impact, based on years of NHD experience.



Source: Student Post-Survey

- In the civic engagement items, we saw a very slight positive change or difference in engagement between NHD and non-NHD students (significant at the $p < .05$ level), or from no years participation in NHD to 3 years participation; changes after 1 year of participation were not significant.

CONCLUSIONS & RECOMMENDATIONS for FURTHER STUDY

Findings from this study suggest that NHD does the things it sets out to do. Students who participate in the program engage in a level of historical research usually reserved for college students or investigative journalists. They are more confident than their peers in their ability to find and use information, and they outperform them on independent performance assessments that ask them to apply those research skills. They also outperform their peers on assessments of writing—what some feel is the “single academic skill most closely associated with college success.”²⁵

There are other indicators that NHD prepares students for the demands of college and careers. Their standardized test scores are often higher, especially in social studies, but also in other academic areas. Many often receive Exemplary or Commended performance ratings. Although it is difficult to credit any single program with student success, there are clear and consistent indications that academic performance improves with successive years of NHD participation.

NHD also prepares students to become part of an informed citizenry. They have a deeper appreciation of history and a more mature perspective on current events—and they begin to think like historians. They exhibit not just cognitive abilities, but also metacognitive skills: they see the big picture in history and the links between past and present events, and their reflections on what historical topics or themes interest them in history suggest a maturity beyond their years.

These accomplishments are not reserved for certain kinds of students: no matter how we analyzed the data—by gender, ethnicity, grade level, and site—NHD students still posted higher scores and higher levels of interest and confidence than their peers.

As we have noted throughout this report, the differences between NHD students and comparison groups were not dramatic. They were, however, very consistent. There is also no mystery about how students develop the skills and dispositions that may give them an edge in college and careers: The program requires a lot of in-depth, sustained work. Teachers set high expectations; judges expect students to be knowledgeable and prepared; students assess their competition and up the bar. The program is not only an excellent way for teachers to meet academic standards, it’s a good way for students to set high standards for themselves.

²⁵ David T. Conley, “Redefining College Readiness” (Eugene, OR: Educational Policy Improvement Center, 2007). Prepared for the Bill & Melinda Gates Foundation.

Further Study

This initial study provides powerful and revealing findings about those who participate in National History Day and, as any useful study, generates as many interesting questions as it answers. Given the breadth and reach of NHD, there are many sub-populations that could be explored, there are a range of outcomes and variables that have yet to be explored, and there are opportunities to explore the program from both the student and the teacher points-of-view. Below we list some of what we identify as the next steps in NHD research, or questions worthy of exploration. There are others that the NHD organization and its larger community may also want to offer, and even others that NHD funders and supporters may want to ask.

- *A study that takes an in-depth look at what teachers do and explores links between classroom practice and student outcomes.* By gathering data on implementation, we could explore questions such as: What impact does the teacher's implementation, or treatment, have on student outcomes, as gauged by their performance on standardized assessments, advancement in competitions, independent performance assessments, and application of skills in other classes? Does, for example, a deliberate focus on writing better equip students to succeed? How do teachers' efforts vary based on grade, student needs, or student population? In what ways does the teacher's level of effort or fidelity to program design affect students' investment of time, self-confidence and satisfaction, or success in the program?
- *A companion study to determine how NHD changes teaching and curriculum.* Documenting what veteran NHD teachers do, or how established NHD programs have defined curriculum or affected practice, we could gather valuable data to answer questions such as: How and how effectively do NHD teachers address—and meet—standards? To what extent do NHD efforts set or change departmental or school-wide, standards-related methods related to, e.g., the degree to which students do self-directed, in-depth research; write full-length essays; or seek out sources beyond required textbooks and resources? Do these efforts and activities affect student and school performance?
- *A study that looks at special populations, such as special education students, at-risk students, English-language learners, or students who do not typically participate in academic programs and competitions.* Questions explored by this study could include: Does NHD improve academic motivation and performance for all students? What impact does NHD participation have on students with special needs? Does participation help these students transition into mainstream academic activities? Does participation lead to changes in ELL designation or produce greater involvement by parents who do not speak English or who are also English Language Learners?

Findings from the three studies outlined above could further define the program's impact and assemble effective, research-based practices.

- *A follow-up study that tracks college or workplace readiness.* We would seek cooperation of faculty and employers, to see if there are differences between NHD students and their peers in students'

writing, reading, synthesizing, note-taking, and research skills; their ability to work in teams; their reliance on multiple sources for research and support; their ability to manage multiple tasks, courses, or assignments; and their perseverance in pursuit of degrees and advancement. We could look across multiple post-secondary occupations, and different higher education choices, including community and junior colleges and online degree programs, as well as four-year colleges and universities. We could also explore whether, through research, we could connect these to 21st Century skills, and measures of them, in any formal ways.

- *A study of communication that includes writing and oral presentations and arguments.* A study such as this would ask: How does the full array of communication skills students gain through NHD compare to the skills other students acquire through routine schoolwork? Do NHD students' skills transfer to other subject areas, and to college entry and AP assessments? In what other intended and unexpected ways do students apply these skills? Do participating students look at communications careers, as well as more academic content areas (e.g., speech communications, media, marketing, acting, visual arts)?
- *A longitudinal study to track new NHD students for 3-4 years.* Although we examined trends based on years of participation, a study of students' development and application of skills, starting with their first exposure to NHD, was beyond the study's scope. A true pre/post analysis, examining performance, competencies, attitudes and dispositions, and academic pursuits as students move from no exposure to NHD to successive years of participation, could ask: What effect does NHD have on students' academic performance and engagement in school? How do skills and attitudes evolve over time? In what ways or to what extent do skills begin to transfer to other subjects? Does NHD affect students' course selections, academic arcs, and long-term goals? With a comparison group, we could see how NHD students compare to their peers in all these domains.

Appendix A: Sample Data Request, Aldine, Texas

Data requested for students at Teague Middle School, Eisenhower HS, Nimitz HS, Comparison School	2007-2008 School Year		2008-2009 School Year		2009-2010 School Year	
	Fall 07	Spring 08	Fall 08	Spring 09	Fall 09	Spring 2010
DEMOGRAPHIC DATA (6 th – 12 th Grades, for each school year) ID, Name, DOB, Ethnicity, Gender, Spec. Ed Status, LEP, Free/Reduced Lunch Status, Grade, Other Special Designation, (e.g., Gifted, IEP, IB program)	X	X	X	X	X	X
TEST SCORE DATA						
*TAKS/TAAS Scores						
Reading (Grades 7, 8, and 9)		X				
Writing (Grade 7)		X				
ELA (Grades 10 and exit level)				X		X
Social Studies (Grades 8, 10, and exit level)				X		X
End of Course Assessments (EOC)						
English I		X		X		X
U.S. History		X		X		X
NAEP (if available)						
Reading (Grade 8)		X		X		
U.S. History (Grades 8 and 12)						X
Civics (Grades 8 and 12)						X
AP Scores (if available)						
History		X		X		X
Government		X		X		X
BEHAVIORAL DATA (Grades 6 th – 12 th ; should be reported individually, not overall)						
Number discipline referrals	X	X	X	X	X	X
Number of suspension incidents	X	X	X	X	X	X
Number days suspended	X	X	X	X	X	X
Number days attended	X	X	X	X	X	X
Number days enrolled	X	X	X	X	X	X
Number move ins	X	X	X	X	X	X
Number move outs	X	X	X	X	X	X
Withdraw dates	X	X	X	X	X	X
OTHER ACADEMIC						
Grades Received By Subject (6 th – 12 th Grades)	X	X	X	X	X	X
Grade Point Average (9 th – 12 th)	X	X	X	X	X	X
AP Courses taken (9 th – 12 th : English, History, Government) (two columns in excel file—enter 1 in column 1 if enrolled; enter student's grade in column 2)	X	X	X	X	X	X
AP scores, if available		X		X		X
Graduation Status (12 th Grade)		X		X		X

If possible, please set the Excel files up the following way:

- Each Excel file with the any of the data listed above should include the student's Aldine ID#, STN, Name, school year, and grade for that school year.
- Create one Excel file for each school year (2006-2007, 2007-2008, etc.). Create a worksheet within this file for each grade, or one worksheet with all grades combined.
- One Excel file (or included as part of demographic data file) for each school year, with all students in grades 6 – 12, that includes ID#, name and the following:
 - Did the student participate in National History Day? (Yes=1; No=0)
 - Did the student participate in National History Day in any previous years? (Yes=1; No=0)
 - If yes, in what grades did they participate? (Actual Grade)

Other School-Level Data, if available:

- Graduation rates: % NHD students vs. non-NHD students
- College admittance rates % NHD students vs. non-NHD students

Research Questions

For comparisons, NHD and non- NHD students will be matched based on gender, ethnicity, need for social services or special classes (LEP, Special Ed, FRL, Gifted), and previous performance on Reading or English Language Arts assessments.

1. Do NHD students differ from their age/grade cohort in academic performance as measured by standardized assessments in reading, language arts, social studies, and other tested subjects? Are differences sustained or deepened with continuing participation?
2. Are there differences in the rates of “commended performance” on standardized assessments between NHD and non-NHD students?
3. Are there differences in rates of passing exit exams *the first time* between NHD and non-NHD students?
4. How many NHD students, compared to non NHD-students, receive the TPSP designation on their high school diplomas?
5. Other areas of impact to explore:
 - a. Changes/trends in attendance (e.g., do attendance rates of students with low or truant records improve when they become involved in NHD? Do rates differ by ethnicity, gender, or LEP status?)
 - b. Changes/trends in school behavior (e.g., do suspensions/behavioral referrals reflect any similar trends? Trends in student mobility?)
 - c. Are there other performance differences, e.g., social studies course grades?
 - d. How do graduation rates differ between NHD and non-NHD students? College admittance rates?
 - e. How many course credits do NHD students earn by high school graduation, on average, and do these numbers differ from courses taken by non NHD-students?
 - f. What percentage of NHD complete advanced courses or AP courses—in science, mathematics, English, as well as in history? Do these percentages vary across student characteristics, including sex, race/ethnicity, etc? What percentage of NHD students take AP exams? How well do they do?

Appendix B: Scoring Rubric for Student Writing

Score	Content (position on issue, support, sense of voice & audience)	Organization (structure, focus, flow)	Word Choice & Fluency (phrasing, vocabulary, sentence variety)	Conventions (spelling, grammar, usage, punctuation)
1	No plausible position is taken on the topic; lacking in examples or evidence. No sense of voice or audience	Writing is disorganized, with little or no focus or coherent argument. Lacks a point of view and central idea	Writing contains fundamental vocabulary mistakes. Severely flawed sentence structure	Grammar and word usage errors interfere with meaning; very poor mechanics and/or punctuation.
2	Position on topic is unclear or extremely limited; inappropriate examples or insufficient evidence. Little sense of voice or audience	Writing is poorly organized; lacks focus; problems with coherence or flow of ideas. Only vaguely suggests a central idea.	Writing shows poor use of language; indicates very limited vocabulary and poor word choice. Frequent problems with sentence structure.	Grammar and word usage mistakes are frequent and interfere with meaning; poor mechanics.
3	Position on topic demonstrates critical thinking skill applied inconsistently; inadequate, redundant, or irrelevant examples and support. Apparent but uneven sense of voice & audience	Organization and focus are limited; demonstrates lapses in coherence or flow of ideas Inconsistent point of view and central idea.	Writing displays developing use of language; some weak vocabulary and poor word selection Lacks sentence variety or has awkward phrasing.	Contains many mistakes in grammar, word usage and mechanics, which sometimes interfere with meaning.
4	Position on topic demonstrates competent critical thinking skill; examples and evidence are generally adequate, relevant, and appropriate. Voice of the writer sometimes comes through.	Writing is generally organized and focused; demonstrates some coherence and attention to the flow of ideas. Central idea is generally clear.	Displays adequate, but inconsistent, use of language; vocabulary used is generally appropriate Good sentence structure and some variety.	Writing contains some mistakes in grammar, word usage and mechanics, but they don't interfere with meaning. On balance, strengths outweigh weaknesses.
5	Position is effectively developed; examples and evidence are adequate, relevant, and appropriate. Voice of the writer comes through often.	Writing is well organized and focused; demonstrates coherence and ideas flow well. Consistent point of view and focus with a central idea throughout.	Displays competent use of language; uses appropriate vocabulary. Good sentence structure and variety.	Writing is generally free of mistakes in grammar, word usage, and mechanics.
6	Position is effectively and insightfully developed; examples and evidence are well-chosen, relevant, and appropriate. Voice of the writer comes through clearly & compellingly.	Writing is well organized, clearly focused, & coherent; ideas flow seamlessly. Maintains a consistent point of view and a clearly presented & supported central idea.	Writing displays skillful use of language; vocabulary is accurate and varied; words are appropriately and skillfully chosen. Meaningful and skilled variety in sentence structure.	Writing is free of most mistakes in grammar, word usage, and mechanics.