Final Evaluation Report for the Teaching American History Project

CENTER FOR RESEARCH ON TEACHING AND LEARNING

NEW YORK UNIVERSITY

Department of Teaching and Learning
82 Washington Square East, Suite 700
New York, NY 10003 | 212 998 5872 | 212 995 3636 fax
www.steinhardt.nyu.edu/teachlearn/crtl
Final Evaluation Report for the Teaching American History Project

Robert Tobias, Director of CRTL
Margary Martin, Research Assistant
Jill Jeffery, Research Assistant
Priscilla Fields, Research Assistant

CRTL External Evaluation Report Series
EE-0206-01

February 2006

Center for Research on Teaching and Learning
Department of Teaching and Learning
The Steinhardt School of Culture, Education, and Human Development
New York University

© Copyright 2006 by the Center for Research on Teaching and Learning
Abstract

This report provides a summative evaluation of the impact of the Teaching American History Project (TAHP) on its participating teachers and students. TAHP was implemented in NYC’s Region 1 in the fall of 2002. The Project worked in collaboration with various contractors, including the New York University Steinhardt School of Education’s Department of Teaching and Learning, the Museum of the City of New York, the Teacher’s Curriculum Institute History Alive program, and Colonial Williamsburg, to provide a number of professional development workshops and summer institutes for participants. A variety of quantitative and qualitative data were collected and analyzed to answer the Project’s evaluation questions.

The evaluation questions addressed the efficacy of the Project in increasing participating teachers’ content knowledge and pedagogical practices in American history. Evaluation of TAHP was also undertaken to determine if a relationship exists between the number of teachers in Region 1 schools who participated in TAHP and the schools’ year-to-year changes in state social studies assessment scores, as well as to assess the Project’s impact on participating teachers’ use of multimedia technology in their day-to-day instruction.

Overall, the Institute assessment data indicate that the program’s Institutes and workshops did have a statistically and educationally meaningful effect upon participating teachers’ content knowledge and pedagogical practices in American history. In fact, some participants of the Project felt that their pedagogical practices had been completely transformed by TAHP. In particular, this Project increased participating teachers’ knowledge of historical leaders and use of primary documents and the arts in the classroom. Dependence on textbooks and lecturing decreased as a result of teacher exposure to TAHP. Participating teachers’ classrooms became more interactive, more dialogic, as a result. Participation in TAHP enabled teachers to increase their students’ engagement with the content of American history as well to make history more relevant to their students’ lives.

Because of methodological limitations, little should be made of the evaluation’s inability to find a relationship between the number of teachers in Region 1 schools who participated in TAHP and the schools’ year-to-year changes in state social studies assessment scores. Also, although the Project did not appear to have a significant impact on participating teachers’ use of multimedia technology in their day-to-day instruction, such issues could be addressed in replications of the Project.

Most participants of TAHP expressed a desire to have the grant extended. Recommendations for improving TAHP include offering wider content knowledge coverage, more instruction on assessment, and more opportunities for teachers to share what they have learned with one another.
I. Background .................................................................1
II. Evaluation Questions ................................................1
III. Evaluation Methods and Procedures ..................................2
   Evaluation Questions 1 and 2: Content Knowledge and Pedagogical Practice ........2
   Evaluation Questions 3 and 5: Perceptions of Pedagogical Competence and
   Use of Multi-Media ..........................................................3
   Evaluation Question 4: School Gains in State Social Studies Assessment Scores .......4
IV. Results ...........................................................................6
   Description of Project Implementation ..................................6
   Project Highlights .............................................................6
   Evaluation Questions 1 and 2: Content Knowledge and Pedagogical Practice ..........7
   Evaluation Questions 3 and 5: Perceptions of Pedagogical Competence and
   Use of Multi-Media ..........................................................11
   Summary of Themes: .........................................................11
   Breakdown of the Narrative Responses by Program Objectives: .....................12
      Objective 1: Content Knowledge ..................................12
      Objective 2: Pedagogical Practice ..................................12
      Objective 3: Leadership Ability ....................................13
      Objective 4: Collaboration ..........................................13
      Objective 5: Use of Multi-Media ...................................13
   Evaluation Question 4: School Gains in State Social Studies Assessment Scores ....15
V. Discussion ......................................................................15
VI. Appendix .......................................................................17
LIST OF TABLES AND FIGURES

Figure 1    Logic Model of the Teaching American History Project ............................................... 5
Table 1    Comparison of Mean Rubric Ratings and Mean Ranks for Pre- and Post-Institute
            Content Assessment Items ........................................................................................................ 8
Table 2    Comparison of Participants’ Reported Percent of Class Time Spent Lecturing
            Before and After Institutes ......................................................................................................... 9
Table 3    Comparison of Participants’ Reported Percent of Planning Resources
            Characterized as Textbooks Before and After Institutes ................................................................. 9
Table 4    Comparison of Participants’ Reported Percent of Class-Work and Homework
            Resources Characterized as Textbooks Before and After Institutes ........................................... 10
Table 5    Comparison of Participants’ Reported Percentage Weight Assigned
            to Short Answer/Multiple-Choice Tests in Assigning Grades Before
            and After Institutes ..................................................................................................................... 10
Table 6    Comparison of Participants’ Reported Use of the Arts and Media in Class
            Instruction Before and After Project Institutes ............................................................................... 11
Table 7    Results from Survey Portion of Questionnaire (N=8) .......................................................... 14
I. BACKGROUND

The Teaching American History Project (TAHP) is a professional development program designed to help participating 4th, 7th, and 8th grade teachers increase their content knowledge of American history and to develop their professional competency in using varied planning, instructional, and assessment strategies for teaching American history. In addition, TAHP aims to increase participating teachers’ ability to integrate the use of multimedia technology in their daily instruction. The project was initiated in fall 2002 in the then NYC Department of Education’s District 10, with support from a federal grant. Subsequently, the program expanded its target population to include teachers in District 9, which was combined with District 10 to form Region 1 as part of the reorganization of the New York City’s public school system following the institution of mayoral control.

The Project worked in collaboration with various contractors, including the New York University Steinhardt School of Education’s Department of Teaching and Learning, the Museum of the City of New York, the Teacher’s Curriculum Institute History Alive program, and Colonial Williamsburg, to provide a number of professional development workshops and summer institutes for participants.

The Center for Research on Teaching and Learning (CRTL) conducted a summative evaluation of the effects of TAHP. CRTL functions within the Department of Teaching and Learning in New York University’s Steinhardt School of Education with the goal of supporting the development and dissemination of effective teacher education and professional development programs through research and evaluation. CRTL used a variety of quantitative and qualitative methods to measure the attainment of the program’s objectives, which were to:

- increase the content knowledge of participating 4th, 7th, and 8th grade teachers;
- enhance the pedagogical practices of participating 4th, 7th, and 8th grade teachers;
- increase the academic achievement of students taught by participants in the program; and
- improve the participating teachers’ success in integrating multimedia technology in their instruction.

II. EVALUATION QUESTIONS

This summative evaluation assessed the impact of TAHP upon teaching and learning of American history in Region 1. The evaluation design was based on the logic model or theory of action for the Project as depicted in Figure 1. The program’s theory of action specifies that by providing teachers with professional development in a variety of formats and venues using a multiplicity of resources, their content knowledge of American history and their pedagogical practice will be enhanced, thereby leading to better teaching and learning, as evidenced by higher scores for the schools of participating teachers on state tests of social studies achievement. Figure 1 indicates the specific outcomes of TAHP’s theory of action that are the focus of this evaluation. Accordingly, the specific evaluation questions are as follows:

1. Do participating teachers demonstrate increased content knowledge in American history after attending TAHP Institutes?
2. Do participating teachers demonstrate enhanced knowledge and understanding of effective pedagogical practices after attending TAHP Institutes?
3. Do participants perceive that TAHP professional development activities enhance their professional competence, i.e., have positive impacts on their pedagogical practices?
4. After controlling for the effects of poverty and the percentage of English language learners, is there a relationship between the number of teachers in Region 1 schools who participated in TAHP and the schools’ year-to-year changes in state social studies assessment scores?
5. To what extent do participating teachers report an increase in their use of multimedia technology in their day-to-day instruction as a result of program participation?

III. EVALUATION METHODS AND PROCEDURES

**Evaluation Questions 1 and 2: Content Knowledge and Pedagogical Practice**

**Subjects:** The subjects for the evaluation of these questions included all teachers who participated in the TAHP Institutes and workshops during the first two years of the Project and completed pre- and post-Institute assessments. The Institutes and workshops were planned and conducted collaboratively by TAHP staff and external contractors. They were conducted over several days with each focused on developing participants’ knowledge of content and innovative pedagogy in the context of specific topics in American history. A total of 70 participating teachers attended the Institutes and completed both the pre- and post-assessments.

**Instruments:** In consultation with Project staff and NYU faculty, CRTL developed Pre- and Post-Institute Assessments (see Appendix). The first four assessment questions were in free-response format and focused on the participants’ knowledge of content in the specific topics of each Institute, including major themes, debatable issues, primary historical documents, and leading historical figures. The next five questions, questions 5 – 9, probed participants’ knowledge and use of effective instructional strategies, including types of classroom activities, types of resources used for instructional planning, class-work, and homework, integration of the arts in instruction, and the use of alternative assessment. Respondents were queried as to the percentages of time spent using a variety of instructional strategies and resources, some of which represented static traditional approaches and others more varied and dynamic approaches that the research literature supports as effective. Promoting increased awareness and use of the latter was a key objective of the program.

**Procedures:** Project staff asked all institute participants to complete the Pre-Institute Assessment at the beginning of each Institute and the Post-Institute Assessment at the conclusion. Each participant was assigned an identification number that was used to pair pre- and post-assessment responses without compromising the participants’ confidentiality. In consultation with Project staff, CRTL developed a scoring rubric for the four content questions. The rubric was based on the historical accuracy, relevance, and depth of knowledge demonstrated by the responses, as well as their completeness. A four-point rubric along a dimension of proficiency was used, with nominal categories as follows: (1) Not Proficient, (2) Partially Proficient, (3) Proficient, and (4) Advanced.
Four raters, three graduate students in social studies education and one Project staff member, were trained to rate the responses. During training, the inter-rater agreement of the scoring process was assessed. The concordance of ratings across all pairs of the four raters was determined throughout the training process. The final levels of agreement were 75% exact agreement and 95% within one score point.

Pairs of raters scored each paper independently with third-party adjudication used to resolve discrepancies of more than one score point. The papers were not identified as pre or post to the raters. All data were entered in an Excel spreadsheet which was imported into SPSS, version 13.0 for analysis.

Data Analysis: For the content knowledge questions, questions 1 – 4, pre- and post-assessment scores were compared statistically using the Wilcoxon Signed Ranks Test, a non-parametric test of statistical significance for ordinal data from related samples. Scores from pre- and post-assessments were ranked in ascending order in a single distribution, and the mean ranks of the pre-assessments and the post-assessments were determined. Only the scores of subjects who completed both pre- and post-assessments were used. The mean pre- and post-score ranks were compared using a z statistic. Z scores with a probability of less than .05 under the assumption of no significant differences were used to reject the null hypothesis of no significant pre- to post-score differences in ranks.

For the pedagogical practices questions, questions 5 – 9, a different analytic method was used. In collaboration with Project staff and based on the research literature, CRTL designated certain choices as static and traditional and others as more dynamic and effective. The mean percentages of these choices in pre- and post-assessments were compared using t-tests for paired data applied at the .05 level of significance.

Evaluation Questions 3 and 5: Perceptions of Pedagogical Competence and Use of Multi-Media

Subjects: The sample consisted of eight participating teachers, five from cohort 1 and three from cohort 2. These teachers participated in a focus group about the efficacy of the Project and its impact upon teaching and learning in their classrooms and responded to a questionnaire during a staff development session on November 2, 2004.

Instruments and Procedures: The participants were first asked to complete a questionnaire that included both Likert-type and open-ended questions. The Likert-type items, based on the goals of the grant, were designed to get a general idea of how the participants felt their participation in the grant affected their knowledge of American history and teaching practice. The open-ended questions were modeled on earlier questionnaire items administered to participants before and after TAHP Institutes. They focused on how the teachers created curriculum units, the resources they used in their classrooms (including multi-media), and pedagogical practices.

The focus group discussion, conducted after participants completed the questionnaires, was based on an interview protocol that prompted participants to assess both strengths of the grant and areas in need of improvement.
Data Analysis: The data from the focus group discussion and the open-ended items in the questionnaire were content analyzed concurrently for recurring themes. Once tallied, the themes were categorized around the objectives of the grant. Frequency distributions of Likert responses were compiled and used to assess the subjects’ perceptions of the attainment of the Project objectives.

EVALUATION QUESTION 4: SCHOOL GAINS IN STATE SOCIAL STUDIES ASSESSMENT SCORES

Subjects: Schools in Region 1 with fifth graders and/or eighth graders participating in the state social studies assessment programs in the school years 2002 – 2003 and 2003 – 2004.

Data: Three sets of data were compiled to address this evaluation question. First, state Social Studies Achievement test scores for grades five and eight were extracted from the New York City Department of Education’s (NYC DOE) Region 1 electronic student information system. The data included aggregate scores for 2002 – 2003 and 2003 – 2004 in two metrics: mean scaled scores and the percent of test takers achieving specific performance levels, Levels 1 – 4 with 3 signifying “proficient” performance and 4 “exceptional” performance. Second, demographic data, including the percent of students eligible for free lunch and the percent of English language learners, were transcribed from Annual School Report cards for 2003 – 2004, which were accessed through the NYC DOE website (www.nycenet.edu/daa). Last, the numbers of teachers from each school who participated in the Project during its first two years were compiled from Project records. These data were used as a school-level measure of the extent of program implementation and faculty exposure to the program for each of the region’s schools. While this measure is only a proxy for the level of a school’s involvement in the program and does not assess the quality of implementation or involvement, it is the best quantitative measure available.

Data Analysis: Regression analysis was used to assess the magnitude of the relationship between the numbers of teacher participants and the size of a school’s change in social studies test scores, controlling for differences in the demographics of a school’s student population. Mean scaled scores and percents of students achieving performance levels 3 and 4 in 2003 – 2004 were separately regressed on the same test scores for 2002 – 2003, demographics, and the number of teacher participants. The first two sets of measures were used as covariates and the last measure was used as the predictor variable in an hierarchical regression model. The significance of the increase in explained variance due to the number of teacher participants was tested to assess the impact of the program upon state social studies test scores. The available data precluded a more sensitive analysis at the individual student or class level.
Figure 1. Logic Model of the Teaching American History Project

**INPUT**
District Training: Content/Pedagogy
- Workshops: Museum of the City of New York (4th Gr) & NYU (7/8th Gr)
- Summer Institutes: Museum of the City of New York (4th Gr) & NYU (7/8th Gr)
- Supplementary: History Alive, Colonial Williamsburg
- Technology
- Mentoring
- Leadership

**COHORTS**
- 4th Gr Teachers
- 7/8th Gr Teachers

**OUTCOME**
Changes in:
- Content Knowledge
- Pedagogical Practice (Target of Evaluation)

**Student Outcomes:**
- Improved student achievement on state mandated tests (Target of Evaluation)

**OUTCOME**
Products:
- Video archive of model lessons
- Website of exemplary American history resources
- CD-Rom virtual field trips
- Disseminate program model at State/National Social Studies conferences

*It is expected that cohort participants will mentor non-cohort teachers in their schools*
IV. RESULTS

The results section begins with a description of the implementation of the Project, including program highlights, followed by the findings from the analyses of data for each of the five summative evaluation questions.

DESCRIPTION OF PROJECT IMPLEMENTATION

TAHP was a collaboration with District 10 and the following partners: New York University, the Museum of the City of New York, Teachers’ Curriculum Institute, and Colonial Williamsburg. The goal of the Project was to improve instruction through the organization of professional development Institutes and workshops for elementary and middle school teachers of American history. These workshops were designed to improve teachers’ content knowledge of American history, increase teachers’ ability to integrate the use of multimedia technology as a tool in the everyday teaching and learning of American history content, and increase students’ achievement in American history.

To ensure that the goals were met, the Project drew upon the intellectual and practical resources of higher education and nonprofit cultural institutions to enhance teachers’ repertoire of historical resources and expand their range of instructional strategies that would also incorporate the use of technology. Based on this intensive experience, teachers were presented with the opportunity to, as lead or master teachers, share the things they learned with their colleagues in professional development workshops and with future cohorts of teachers involved in the Project.

The Project activities were structured into four major components: four full day professional development workshops at the Museum of the City of New York (grade 4 teachers) or New York University (grade 7-8 teachers), four study group sessions on Saturdays, a summer Institute at New York University or the Museum of the City of New York, and a Teachers’ Curriculum Institute training session. Several teachers each summer were invited to attend a full week Institute at Colonial Williamsburg. According to an interview conducted with teachers, they all agreed that this experience increased their content base of knowledge and improved their instructional practice.

PROJECT HIGHLIGHTS

Colonial Williamsburg: Twenty-one teachers attended the Colonial Williamsburg Summer Institute over the last three years. Each participant served a leadership role in that he/she conducted a workshop and/or produced a lesson for posting on the Project’s website based on the numerous workshops attended that week. Following the Institute, several teachers from the three cohorts engaged in electronic field trip experiences with their own students. These electronic field trips were developed by Colonial Williamsburg on various historical topics. Teachers shared these materials with their colleagues. Project participants learned innovative strategies for infusing technology into the elementary and middle school curriculum by also attending workshops that Colonial Williamsburg provided during the school year.

New York University: Participants in the Project benefited from the Institutes and workshops that were offered by New York University. The lectures on various topics in American history
were interactive; teachers were able to transfer the historical content and pedagogical strategies they had acquired as evidenced by informal classroom observations by the Project Director in the second and third year of the Project. Teachers, on videotaped conversations and on written evaluations, reflected on the knowledge gained and how they would teach different eras in history differently. Their bimonthly experience was also measured by pre- and post-assessments completed during the workshops and the lessons and units of study that were submitted to the Project Director at the end of the school year. Regional officials indicated to the evaluators their intent to upload onto its TAHP website the videotaped lessons and units of study in 2006.

**Museum of the City of New York**: Elementary school teachers primarily participated in workshops at the Museum of the City of New York. This cultural institution has numerous artifacts and exhibits that are relevant to the study of local history. Teachers, many of whom had limited historical knowledge about major events in American history, indicated that they were able to glean historical information and skills for study from the period in which various Native American tribes inhabited the colony. Engaging in interactive tours around various neighborhoods showed fourth grade teachers how to involve their own students in studying how communities develop and change over time. Teachers were also trained to help their colleagues develop local history lessons, projects, and units of study.

**Teachers’ Curriculum Institute**: Through this multi-day training Institute, elementary and middle school teachers examined innovative ways to bring history alive in their classrooms. They were introduced to historical periods in American history and were challenged to consider how students could create illustrated timelines, participate in experiential exercises that assisted them in understanding the various perspectives that existed during the revolutionary periods, and complete project-related assignments that encouraged students to think creatively and draw upon their prior knowledge. Observations of participants’ lessons by the Project Director indicated that many of the teachers’ lessons reflected the practices that were the focus of the Institute. The Project Director observed that participating teachers’ students’ perceptions of history had changed. Students in the most challenging schools had a different attitude toward the subject.

**EVALUATION QUESTIONS 1 AND 2: CONTENT KNOWLEDGE AND PEDAGOGICAL PRACTICE**

**Content Knowledge**: Both quantitative and qualitative methods were used to assess the Project’s effectiveness in increasing participating teachers’ content knowledge of American history and knowledge and understanding of effective pedagogical practices. Only the responses of the 70 participants who completed both pre- and post-Institute assessments were used to evaluate questions 1 and 2.

Participants’ responses to the first four questions of the Institute assessments (see Appendix) were used to answer Evaluation Question 1. These four questions probed participants’ knowledge of various events in American history. Question one asked participants to list three themes they stress when teaching a particular topic. Question two asked participants to identify a resolved proposition that they would use in class to engage their students in educational debate. Question three asked participants to name two primary historical documents that they would use to teach a particular event in history and why they would choose such documents. Question four asked participants to name two historical figures they would use to teach a particular event in history and why they would choose such figures. Participants’ responses were rated using a four-
point rubric, as mentioned in the procedures section of this report. Table 1 below displays the pre- and post-institute assessment mean rubric scores and mean ranks for the four content knowledge questions, as well as the total content assessment scores. The statistical significance of differences in pre- and post-scores was tested through the application of Wicoxon Signed Ranks tests to the data for each item and the total content assessment scores. Post-test scores were significantly larger than pre-test scores for the total content assessment score, mean pre-test score = 9.02 and mean post-test score = 10.12, $z = -2.53$, $df = 59$, $p = .011$. Deeper inspection of the data show statistically significant gains in knowledge of historical leaders, mean pre-test score = 1.97 and mean post-test score = 2.37, $z = -2.762$, $df = 59$, $p = .006$, and use of primary documents, mean pre-test score = 2.13 and mean post-test score = 2.45, $z = -2.276$, $df = 59$, $p = .023$. The effect sizes for these scores indicated small, educationally meaningful gains in these areas. The gain in knowledge of major historical themes was not statistically significant but did indicate a small, educationally meaningful effect for the Institutes. The gain in formulation of debate topics was neither statistically significant nor educationally meaningful.

Table 1
Comparison of Mean Rubric Ratings and Mean Ranks for Pre- and Post-Institute Content Assessment Items

<table>
<thead>
<tr>
<th>Assessment Scores</th>
<th>N</th>
<th>Mean Rating (Scale: 1-4)</th>
<th>SD</th>
<th>Mean Rank</th>
<th>Z</th>
<th>p value</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge of Major Historical Themes</td>
<td>Pre</td>
<td>70</td>
<td>2.50</td>
<td>0.88</td>
<td>19.35</td>
<td>-1.804</td>
<td>0.071</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>70</td>
<td>2.73</td>
<td>0.74</td>
<td>19.58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Formulation of Debate Topics</td>
<td>Pre</td>
<td>60</td>
<td>2.40</td>
<td>1.01</td>
<td>16.82</td>
<td>-0.504</td>
<td>0.614</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>60</td>
<td>2.50</td>
<td>0.95</td>
<td>19.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use of Primary Documents</td>
<td>Pre</td>
<td>60</td>
<td>2.13</td>
<td>1.00</td>
<td>13.32</td>
<td>-2.276</td>
<td>0.023</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>60</td>
<td>2.45</td>
<td>0.87</td>
<td>18.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knowledge of Historical Leaders</td>
<td>Pre</td>
<td>60</td>
<td>1.97</td>
<td>0.84</td>
<td>15.60</td>
<td>-2.762</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>60</td>
<td>2.37</td>
<td>0.86</td>
<td>18.96</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Content Assessment Score</td>
<td>Pre</td>
<td>60</td>
<td>9.02</td>
<td>2.75</td>
<td>19.06</td>
<td>-2.533</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>60</td>
<td>10.12</td>
<td>2.38</td>
<td>27.77</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Scale Values: (1) Not Proficient; (2) Partially Proficient; (3) Proficient; (4) Advanced.

Overall, the Institute assessment data indicate that the program’s Institutes and workshops did have a small but educationally meaningful effect upon participating teachers’ content knowledge in American history. These findings were supported by qualitative data reported later in the results section.

**Pedagogical Knowledge:** Participants’ responses to questions five - nine on the Institute assessment were used to answer Evaluation Question 2. Assessment question five asked participants to indicate the percentages of class time they typically spend using the following instructional methods: lecture, whole class discussion or debate, small group discussion, and other. An analysis of participants’ responses is presented in Table 2 below. A goal of the program was to
develop the participating teachers’ awareness of the importance of varying instructional methodologies, especially reducing the amount of time spent lecturing. To evaluate the effects of the Institutes and workshops on participants’ instructional methods, they were asked to report the percent of time they spent using various methods before and after the Institutes. Participants reported spending an average of 61.2% of class time lecturing before the Institutes compared to 52.8% following the Institutes. The reduction in mean percent of time lecturing was statistically significant, \( t = -4.314, df = 33, p = .001 \). The 8.4 percentage point decline in lecturing indicates a moderate, educationally meaningful effect upon instruction.

### Table 2

Comparison of Participants’ Reported Percent of Class Time Spent Lecturing Before and After Institutes

<table>
<thead>
<tr>
<th>Assessment</th>
<th>N Teachers</th>
<th>Mean Percent Class Lecture</th>
<th>Standard Deviation</th>
<th>Mean Difference</th>
<th>Standard Error</th>
<th>Degrees of Freedom</th>
<th>t-value</th>
<th>P Value</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre</td>
<td>34</td>
<td>61.18</td>
<td>14.72</td>
<td>-8.38</td>
<td>1.94</td>
<td>33</td>
<td>-4.314</td>
<td>.001</td>
<td>-0.57</td>
</tr>
<tr>
<td>Post</td>
<td>34</td>
<td>52.79</td>
<td>18.27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Assessment questions six and seven asked participants to consider the resources they typically use in planning lessons and in class-work and homework. What percentage could be classified as primary sources, textbooks, web-based resources, history books (not texts), and other? An analysis of participants’ responses is presented in Tables 3 and 4 below. A goal of the program is to broaden the array of resources used in planning, class-work, and homework and, in the process, reduce reliance on textbooks. As seen in Tables 3 and 4, respectively, participating teachers showed a statistically meaningful and educationally meaningful reduction in the use of textbooks, and a concomitant increase in the use of other resources, for both planning and class-work and homework. Mean percent of reported use of textbooks for planning declined from 32.6% before the Institutes to 24.4% after the Institutes, \( t = -2.923, df = 32, p = .006 \). The reduction in textbook use for class-work and homework was similarly large, from 34.7% before to 24.2% after the Institutes, \( t = -2.747, df = 32, p = .01 \).

### Table 3

Comparison of Participants’ Reported Percent of Planning Resources Characterized as Textbooks Before and After Institutes

<table>
<thead>
<tr>
<th>Assessment</th>
<th>N Teachers</th>
<th>Mean Percent Textbook Use</th>
<th>Standard Deviation</th>
<th>Mean Difference</th>
<th>Standard Error</th>
<th>Degrees of Freedom</th>
<th>t-value</th>
<th>P Value</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre</td>
<td>33</td>
<td>32.58</td>
<td>19.04</td>
<td>-8.21</td>
<td>2.81</td>
<td>32</td>
<td>-2.923</td>
<td>0.006</td>
<td>-0.43</td>
</tr>
<tr>
<td>Post</td>
<td>33</td>
<td>24.36</td>
<td>13.89</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 4
Comparison of Participants’ Reported Percent of Class-Work and Homework Resources Characterized as Textbooks Before and After Institutes

<table>
<thead>
<tr>
<th>Assessment</th>
<th>N Teachers</th>
<th>Mean Percent Textbook Use</th>
<th>Standard Deviation</th>
<th>Mean Difference</th>
<th>Standard Error</th>
<th>Degrees of Freedom</th>
<th>t-value</th>
<th>P Value</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre</td>
<td>33</td>
<td>34.70</td>
<td>23.25</td>
<td>-10.48</td>
<td>3.82</td>
<td>32</td>
<td>-2.747</td>
<td>0.01</td>
<td>-0.45</td>
</tr>
<tr>
<td>Post</td>
<td>33</td>
<td>24.21</td>
<td>16.65</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Institute assessment question nine asked participants to consider how they determine student report card grades. What approximate percentage weight do participants give to short answer/multiple-choice tests, essays, reports and projects, class participation, presentations, and other? An analysis of participants’ responses is presented in Table 5 below. Contrary to the goals of the program, participants did not report a statistically significant or educationally meaningful change in the percentage weight they assigned to short-answer/multiple-choice tests in the determination of grades after participating in the Institutes.

Table 5
Comparison of Participants’ Reported Percentage Weight Assigned to Short Answer/Multiple-Choice Tests in Assigning Grades Before and After Institutes

<table>
<thead>
<tr>
<th>Assessment</th>
<th>N Teachers</th>
<th>Mean Percent Short Answer</th>
<th>Standard Deviation</th>
<th>Mean Difference</th>
<th>Standard Error</th>
<th>Degrees of Freedom</th>
<th>t-value</th>
<th>P Value</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre</td>
<td>29</td>
<td>23.33</td>
<td>12.07</td>
<td>0.62</td>
<td>1.78</td>
<td>28</td>
<td>-0.348</td>
<td>0.73</td>
<td>-0.05</td>
</tr>
<tr>
<td>Post</td>
<td>29</td>
<td>22.71</td>
<td>10.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Finally, Table 6 below shows an analysis of participants’ responses to Institute assessment question eight: “How often do you use art, photographs, music and plays in classroom instruction?” Participants responded to the question using a 5-point Likert-type scale ranging from (1) Never to (5) Very Often. A goal of the Institutes was to promote awareness of the efficacy of the arts and media to support classroom instruction and to provide concrete examples of ways to incorporate them in the classroom. As seen in Table 6, the Institutes appear to have been successful in this endeavor. Participants reported a statistically significant, as determined by the application of a Wilcoxon Signed Ranks test to the data, and highly meaningful increase in their use of the arts in classroom instruction before and after the Institutes. Usage increased from a mean of 3.13, sometimes, before the Institutes to 4.26, often, after, $z = 2.84$, df = 69, $p = .005$. 
Table 6
Comparison of Participants’ Reported Use of the Arts and Media in Class Instruction Before and After Project Institutes

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean Rating (Scale: 1-4)</th>
<th>SD</th>
<th>Mean Rank</th>
<th>Z</th>
<th>p value</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre</td>
<td>70</td>
<td>3.13</td>
<td>0.92</td>
<td>6.00</td>
<td>2.840</td>
<td>0.005</td>
<td>1.23</td>
</tr>
<tr>
<td>Post</td>
<td>70</td>
<td>4.26</td>
<td>1.13</td>
<td>6.55</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Scale Values: (1) Never (2) Not Often (3) Sometimes (4) Often (5) Very Often

EVALUATION QUESTIONS 3 AND 5: PERCEPTIONS OF PEDAGOGICAL COMPETENCE AND USE OF MULTI-MEDIA

Overall, the assessment data indicate that the program successfully achieved its objective of improving the pedagogical knowledge and practices of participants. Further evidence for this finding is provided by the qualitative data reported below.

Perceptions of the impact of the program on participants’ pedagogical competence were obtained through the administration of a follow-up questionnaire and the convening of a focus group at a program workshop conducted at the Museum of the City of New York on November 2, 2004. The participants were first asked to complete the questionnaire that included both Likert-type and open-ended questions. The questions, based on the goals of the grant, were designed to get a general idea of how the participants felt their participation in the grant affected their knowledge and teaching practice. The results of the Likert questions are displayed in Table 7 at the end of this section. The narrative responses of participants to the questionnaire and the themes that emerged in the focus group discussion were content analyzed and are summarized below.

Summary of Themes:

- Overall, the respondents were highly positive in their comments about the Project. Even when prompted to comment on problems with the program and suggestions for improving it, the responses were often stated in terms of a desire to extend the grant. Participants wanted more content, more instruction on assessment, more opportunities to share what they learned.

- The professional development appears to have transformed the participants both in terms of content knowledge and pedagogy. Numerous participants referred to the transformative power of incorporating primary documents into their curricula. This led them to decrease their dependence on textbooks; it helped them engage students in the content; and, perhaps most importantly, it engaged the teachers themselves in the content more fully. In addition, teachers repeatedly mentioned the many instructional strategies to which they were introduced that helped them integrate new content and resources into their curricula.

- Another theme that emerged from the focus group discussion was the teachers’ desire to transform their schools. In one case, a participant sought out and obtained the opportunity
to share her new knowledge with her colleagues in professional development sessions. Others advanced arguments for integrating social studies content with literacy instruction on their campuses. Many expressed frustrations with school administrators who relegated social studies to secondary status as a result of pressure to raise ELA test scores. A few mentioned a desire for help to procure heavier investment in social studies from their school administrators.

**Breakdown of the Narrative Responses by Program Objectives:** After determining the major themes, the content analysis proceeded by breaking the narratives down by the key objectives of the program. This breakdown is presented below, with the narrative comments organized by what went well and suggestions for improvement.

**Objective 1: Content Knowledge**

What went well

A. Themes
   - Increased teachers’ content knowledge
   - Received large number of beneficial resources

B. Activities (where content was received)
   - The Summer Institute
   - Colonial Williamsburg
   - History Alive

C. Resource Use
   - Don’t use textbooks as much, if at all
   - Increased use of primary documents (everyone mentioned), i.e. Dear Mrs. Roosevelt letters, Anderson and Reconstruction letters
   - History Alive
   - Artifacts
   - Visual media (film, cartoons, photos)

Suggestions for improvement
   - Feel they have gaps in content knowledge/Would like wider content coverage
   - Few historian presenters of color
   - More multicultural history/perspectives
   - Planning—deciding what to cut (overwhelming amount of resources)
   - Would like field trips to NYC historical sites/museums

**Objective 2: Pedagogical Practice**

What went well

A. Themes
   - Teachers felt their pedagogical practices had been completely transformed.
   - They are now able to engage students more in the content and make history relevant to their students.

B. Activities
   - Institutes
   - Workshops
   - Colonial Williamsburg
C. Examples

- Theatre activities: i.e. drama, Tableau, popcorn
- Fishbowl
- Debate
- Using artifacts

Suggestions for improvement

- No assessment instruction. Need to know how to assess student performances, especially in theatre activities
- Need more explicit teaching of how to link literacy to history
- More hands-on/interactive instruction (i.e. more theater)
- Have video clips in classrooms showing example pedagogy in practice

Objective 3: Leadership Ability

What went well

- Teachers have worked to share with teachers in schools and take leadership roles at their schools with varying levels of success.
- Examples individuals gave of successes: sharing materials and ideas with colleagues, helping new teachers, showing other literacy teachers how to link literacy to social studies

Suggestions for improvement

- There have been some major constraints on teachers’ ability to take on leadership roles in their schools. Most often noted is the diminished priority of the teaching of social studies due to ELA testing pressures.
- Help them to convince administrators of the importance of the subject

Objective 4: Collaboration

What went well

- Sharing information and ideas with others in their cohort during the study group sessions
- Feel they are developing a network of teachers

Suggestions for improvement

- Not enough time made for sharing, either within their schools or outside of their schools
- More opportunities for networking are needed

Objective 5: Use of Multi-Media

- There is an absence of narrative discussion of this objective in both the questionnaire and focus group responses. This omission suggests that the topic did not receive the attention that was initially intended.
- This interpretation of the data is supported by the responses to Likert Question 9, *I can better incorporate technology into my teaching* (see Table 7.) Six of the eight respondents answered Moderately Disagree to this question.
### Table 7
Results from Survey Portion of Questionnaire (N=8)

<table>
<thead>
<tr>
<th>Part I: Please indicate the extent to which you agree or disagree with each statement below by circling the appropriate numeral to the right of each statement.</th>
<th>1 Strongly Disagree</th>
<th>2 Moderately Disagree</th>
<th>3 Moderately Agree</th>
<th>4 Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. My knowledge of U.S. history has increased as a result of my participation in the Project.</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>2. I now have a larger range of resources on which to develop new curriculum compared to before my participation in the Project.</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>3. My repertoire of instructional methods has expanded as a result of my participation in the Project.</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>4. The Project helped me to build a sustained, collaborative network with the other teachers in the project.</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>5. I now have more ability to affect curriculum development at my school.</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>6. I share what I’ve learned with my colleagues at my school.</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>7. I have become a mentor to other history teachers at school.</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>8. I have become (or plan to become) a mentor to newer participants in the Project.</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>9. I can better incorporate technology into my teaching.</td>
<td>0</td>
<td>6</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>10. I am better able to engage students in the study of history now than before my participation in the Project.</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>11. I am better able to promote student mastery of the state/city learning standards in US history now than before my participation in the Project.</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>12. I am better able to promote deeper student understanding of U.S. history now than before my participation in the Project.</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>6</td>
</tr>
</tbody>
</table>
**EVALUATION QUESTION 4: SCHOOL GAINS IN STATE SOCIAL STUDIES ASSESSMENT SCORES**

Linking the effects of professional development to gains in student achievement test scores is a difficult proposition even when evaluators have wide latitude in controlling the conditions of the evaluation. Short of randomized trials, the most sensitive evaluation models employ pre- and post-testing of students who were directly instructed by the teacher participants with comparisons of gains to equivalent students instructed by non-participating teachers with similar background and experience. In this study, CRTL was limited to (1) employing extant data from state social studies tests administered to students annually in grades 5 and 8; (2) assessing gains at the school level rather than for individual students taught by Project teachers; and (3) using the variability of the numbers of trained teachers in each school as a natural control. Admittedly, the methodological shortcomings of this design severely reduce the prospects of detecting Project effects upon test scores. Nevertheless, given the importance of attempting to assess the ultimate objective of any professional development, gains in student achievement, the evaluators proceeded with the analysis.

Regression analysis was used to try to assess the relationship between the numbers of teachers in a school who participated in the Project and gains in the school’s performance on the state social studies assessment at grades 5 and 8. The logic is that the more trained teachers in a school, the greater the overall impact upon teaching and learning in the school, and consequently the greater the improvement in student learning as measured by the test. To partially control for the effects of student demographics, differences among schools in the percents of students eligible for free lunch and English language learners were partialed out of the analysis through co-variation. After controlling for the demographic variables, the regression analysis showed no statistically significant or educationally meaningful effect of the number of trained teachers on gains in state social studies test scores for the schools in Region 1. For the reasons explained above, this finding should not be given much weight as evidence in the evaluation of the efficacy of the program.

**V. DISCUSSION**

CRTL’s evaluation of Region 1’s TAHP found considerable evidence of the efficacy of the program in improving the teaching and learning of American history in the region’s schools. Overall, the Project assessment data indicate that the program’s Institutes and workshops did have a small but educationally meaningful effect upon participating teachers’ content knowledge in American history. In particular, the Project increased participating teachers’ knowledge of historical leaders and use of primary documents in planning and delivering lessons. The Project had less of an educationally meaningful impact on participating teachers’ knowledge of historical themes and no educationally meaningful impact on their formulation of debate topics.

The Project’s Institutes enhanced participating teachers’ knowledge and understanding of effective pedagogical practices as well as their content knowledge in American history. As a result of the PD, participating teachers reduced their use of lecturing in their classrooms. After attending the Institutes, participating teachers used a wider variety of instructional methods to engage their students in their curricula. Classroom instruction became less didactic and more dialogic, more interactive. Positive changes in pedagogy can be attributed to participating teachers’ exposure to a wide range of instructional resources and an increase in their repertoire of instructional methods. Exposure to a wide range of instructional resources helped reduce participating teachers’ dependence on textbooks in planning class-work and homework assignments—a goal of the Project. A concomitant increase in the use of other resources followed the reduction on the dependence of textbooks. The Project was also successful in its aim to increase participating teachers’ use of art, photographs,
music and plays in classroom instruction in American history. Use of such media increased student engagement and understanding of American history.

The TAHP also intended to impact the way that participating teachers determined student report card grades. The Project hoped to lessen participating teachers’ use of short-answer/multiple-choice tests and increase their use of other assessment means such as essays, reports and projects, and presentations. Contrary to the goals of the program, however, after attending the Institutes, participants did not report an educationally meaningful change in the percentage weight they assigned to short-answer/multiple-choice tests in the determination of final grades. These findings indicate a need to provide more PD in assessment.

Due to methodological shortcomings of this evaluation design, the prospects of detecting Project effects upon state social studies test scores were severely reduced. Accordingly, little should be made of the fact the regression analysis showed that the Project had no meaningful effect of the number of trained teachers on gains in state social studies scores for the schools of Region 1. The Project also had little significant impact on participating teachers’ use of multi-media technology in their day-to-day instruction.

This discussion could not be completed without offering recommendations to be considered for future replications of the TAHP. Results of the focus group discussion on November 2, 2004, showed that overall, respondents were highly positive in their comments about the Project; their responses showed their desire to have the grant extended. Some suggestions for improvement of the Project include offering wider content coverage (more multicultural history and perspectives), more instruction on assessment, and more opportunities for teachers to share what they have learned with one another at their schools. Participating teachers also expressed a desire to attend more field trips to NYC historical sites and museums. Replications of this Project could also address participating teachers’ desire to obtain heavier investment in social studies from their school administrators. Perhaps the link between literacy and social studies could be strengthened. Finally, more emphasis and time could be devoted to the development of a website of exemplary American history resources and a video archive of model lessons. The preceding recommendations could significantly strengthen the efficacy of the TAHP.
APPENDIX

Sample Assessments

PRE-INSTITUTE ASSESSMENT
THE AMERICAN REVOLUTION

Your answers to the following questions will help us to evaluate the effectiveness of our program and to improve future sessions. Your responses will be held in strict confidence and will not be revealed to any of your supervisors or colleagues. These data will be aggregated for analysis and will not be used to evaluate you or any other individuals participating in the program. We appreciate your cooperation.

Name: ______________________________
School:______________________________
Undergraduate Major: _____________________________
Total Number of Courses in American History: __________
License: Type (Check one): PPT ____  CPT _____ Certified _____
Subject(s) _____________________________
Number of Years Teaching: ______________
Highest Degree Earned: ______________________

1. When you teach about the American Revolution, what three themes do you stress?
   1. __________________________________________________________________________
   2. __________________________________________________________________________
   3. __________________________________________________________________________

2. If you were to organize a classroom debate on the American Revolution, what proposition (resolved) would you use to engage your students?
   __________________________________________________________________________
____________________________________________________________________________
3. Name two primary historical documents you would use to teach the American Revolution and explain why you would choose them.

________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

4. Name two historical leaders or figures you would use to teach the American Revolution and explain why you would choose them.

________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

5. Indicate the percentages of your class time on the American Revolution that you typically spend using the following instructional methods:

   Lecture       ________%
   Whole Class Discussion or Debate ________%
   Small Group Discussion ________%
   Other (Please Specify) ________%

                  ________   Total    ________% 100

6. Of the resources that you typically use in planning lessons on the American Revolution, what percentages can be classified as:

   Primary Sources ________%
   Textbooks ________%
   Web-Based Resources ________%
   History Books (Not Texts) ________%
   Other (Please Specify) ________%

                  ________   Total    ________% 100
7. Of the resources that you typically use in class-work and homework assignments on the American Revolution, what percentages can be classified as:

<table>
<thead>
<tr>
<th>Resource</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary Sources</td>
<td>________%</td>
</tr>
<tr>
<td>Textbooks</td>
<td>________%</td>
</tr>
<tr>
<td>Web-Based Resources</td>
<td>________%</td>
</tr>
<tr>
<td>History Books (Not Texts)</td>
<td>________%</td>
</tr>
<tr>
<td>Other (Please Specify)</td>
<td>________%</td>
</tr>
<tr>
<td>Total</td>
<td>100%</td>
</tr>
</tbody>
</table>

8. How often do you use art, photographs, music and plays in classroom instruction on the American Revolution? (Check One)

- Very Often
- Often
- Sometimes
- Not Often
- Never

9. In determining student report card grades, what is the approximate percentage weight that you give to:

<table>
<thead>
<tr>
<th>Component</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short-Answer/Multiple-Choice Tests</td>
<td>______%</td>
</tr>
<tr>
<td>Essays</td>
<td>______%</td>
</tr>
<tr>
<td>Reports and Projects</td>
<td>______%</td>
</tr>
<tr>
<td>Class Participation</td>
<td>______%</td>
</tr>
<tr>
<td>Presentations</td>
<td>______%</td>
</tr>
<tr>
<td>Other (Please Specify)</td>
<td>______%</td>
</tr>
<tr>
<td>Total</td>
<td>100%</td>
</tr>
</tbody>
</table>