

**Baseline TAH 11th Grade (N=13) NAEP Teacher Content Assessment
Summary of Results – December 16, 2008**

Thirteen high school teachers participating in the Oakland Unified School District’s Teaching American History (TAH) project successfully completed a baseline administration of the NAEP Teacher Content Assessment. Project staff created the assessment from a bank of 358 U.S. history/social science questions made available online from the National Assessment for Educational Progress (NAEP). The assessment was administered by an evaluator from the Center for Evaluation and Research (CER) at Mills College on October 28, 2008. Teachers were given as much time as needed to complete the assessment, which were then taken to CER for scoring.

The scoring system was developed by CER and implemented by a member of the CER evaluation team. The 20-question assessment included 10 multiple-choice and 10 short answer questions with a total value of 40 points. The assessment was developed to address one of two TAH GPRA (Government Performance and Results Act) Performance Measures. GPRA Performance Measure 1.1 focuses on student academic achievement while GPRA Performance Measure 1.2 hopes to demonstrate “an increased understanding of American history through the use of nationally validated tests of American history that can be directly linked to their participation in the Teaching American History program.” Results from the baseline administration of the NAEP Teacher Content Assessment are presented in the following tables:

Table 1: Baseline Grade 11 NAEP Results by Mean, Median, Minimum and Maximum Score N=13

	Sum of MC Questions (10 points possible)	Sum of Short Answer Questions (30 points possible)	Total Score (40 points possible)	Total Percent Correct
Mean	9.5	25.5	35	87.5
Median	10	26	36	90.0
Minimum	7	20	30	75.0
Maximum	10	29	39	97.5

Table 1 shows that the average score on the assessment was 35 out of 40 points and the median score was 36 points for the 11th grade group. The lowest score recorded was 30 points while the highest score was 39 points (97.5%). The total scores from all 13 teachers are presented in Table 2.

Table 2: Baseline Grade 11 NAEP Results by Total Score (40 points possible)

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 30	1	7.7	7.7	7.7
31	1	7.7	7.7	15.4
32	1	7.7	7.7	23.1
33	1	7.7	7.7	30.8
34	1	7.7	7.7	38.5
35	1	7.7	7.7	46.2
36	3	23.1	23.1	69.2
37	1	7.7	7.7	76.9
38	2	15.4	15.4	92.3
39	1	7.7	7.7	100.0
Total	13	100.0	100.0	

The NAEP assessment consisted of 10 multiple-choice questions that were valued at one point each. Teachers averaged 9.5 of 10 correct (95%) while the median score was 10 correct. The multiple-choice scores from all the teachers are presented in Table 3 below:

Table 3: Baseline Grade 11 Results Showing Amount of Multiple-Choice Questions Correct (10 Points Possible)

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 7	1	7.7	7.7	7.7
8	1	7.7	7.7	15.4
9	1	7.7	7.7	23.1
10	10	76.9	76.9	100.0
Total	13	100.0	100.0	

Ten teachers got all 10 multiple-choice questions correct. The low score was 7 of 10 (70%). Of the 10 multiple choice questions, 5 questions were answered correctly by all 13 teachers. Those questions are presented in Table 4:

Table 4: One-Point Multiple-Choice Questions that were Answered Correctly by All Teachers N=13

One-point Multiple-Choice Question	1 = Correct	0 = Incorrect
1. Question 8: Immigration sources: Show secondary source	100.0%	0.0%
2. Question 9: Immigration sources: Show attitudes	100.0%	0.0%
3. Question 10: Explain purpose of the Progressive Movement 1890-1920	100.0%	0.0%
4. Question 11: Hamilton's criticism of Articles	100.0%	0.0%
5. Question 14: Great Society assumptions	100.0%	0.0%

Five of the multiple-choice questions received incorrect responses. Four questions received one incorrect response and one received three incorrect responses. The five questions that garnered incorrect responses are presented in Table 5:

Table 5: One-Point Multiple-Choice Questions that were Answered Incorrectly by One or More Teachers N=13

One-point Multiple-Choice Question	1 = Correct	0 = Incorrect
1. Question 6: 1970 U.S. economy affected by	76.9%	23.1%
2. Question 12: Identify message of Bryan's "Cross of Gold" speech	92.3%	7.7%
3. Question 13: Identify groups in favor of the gold standard	92.3%	7.7%
4. Question 15: Depression: long criticized FDR for	92.3%	7.7%
5. Question 20: Identify early American labor union	92.3%	7.7%

The NAEP assessment contained 10 short answer questions that were valued at three points each for a total of 30 points. The project used a three-point scoring guide that was based on NAEP guidelines to score the short answer questions. The scoring guide used four rating levels that ranged from 0 (lowest) to 3 (highest): The rating levels enabled respondents to earn partial credit for incomplete or partially correct scores. The rating levels included: 0) Incomplete, missing, or wrong; 1) Only partially correct, complete and substantial; 2) Mostly correct and complete; and 3) Complete, accurate and substantial. Table 6 on the following page displays the scores from all 13 teachers for the 10 short answer questions.

Table 6: Baseline Grade 11 Results Showing Amount of Short Response Questions Correct (30 Points Possible)

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 20	1	7.7	7.7	7.7
21	1	7.7	7.7	15.4
22	1	7.7	7.7	23.1
24	1	7.7	7.7	30.8
26	3	23.1	23.1	53.8
27	3	23.1	23.1	76.9
28	2	15.4	15.4	92.3
29	1	7.7	7.7	100.0
Total	13	100.0	100.0	

One 11th grade teacher scored 29 of 30 possible points (96.7%) for the highest score, while the lowest score was 20 (66.7%). Teachers averaged 25.5 points out of 30 and the median score was 26 points. Of the ten short answer questions, seven were answered correctly by at least 75% of the teachers. All ten questions are presented from easiest to hardest in Table 7 below:

Table 7: Three-Point Questions Listed in Order from Easiest to Hardest N=13

Three-point Short Answer Question	3 = Complete, accurate & substantial	2 = Mostly correct & complete	1 = Only partially correct, complete & substantial	0 = Incomplete, missing or wrong
1. Question 7: Factors in the growth of suburbs	100.0%	0.0%	0.0%	0.0%
2. Question 5: Reasons for decreased immigration	84.6%	7.7%	7.7%	0.0%
3. Question 16: Depression: Which cause most important	84.6%	7.7%	0.0%	7.7%
4. Question 4: Explain sharecropping's rise	76.9%	23.1%	0.0%	0.0%
5. Question 3: Identify two issues that caused the civil war	76.9%	15.4%	7.7%	0.0%
6. Question 17: Two stats that show severity and what effect spending	76.9%	15.4%	7.7%	0.0%
7. Question 18: Depression: Compare 1933 and 1937 inaugural addresses	76.9%	15.4%	7.7%	0.0%
8. Question 19: Depression: What poem and song show	38.5%	53.8%	7.7%	0.0%
9. Question 1: Use spending tables to show changes	30.8%	53.8%	15.4%	0.0%
0. Question 2: Explain reaction of factory owners to the formation of unions	30.8%	15.4%	53.8%	0.0%

All teachers received three points for correctly answering Question 7. Question 7 asked respondents to explain how either automobiles, the highway system, or tax deductions for mortgage interest contributed to the growth of suburbs. To receive three points, respondents needed to explain the link between one of the factors listed and suburbanization, citing specifics or elaborating on the explanation. One respondent who chose automobiles wrote: "Automobile ownership had a major impact on the growth of suburbs – it was now easier in many parts of the country (California for instance) to live further away from the city where most jobs were located." Another teacher who chose

tax deductions wrote, "Tax deductions for the interest on mortgages can make it cheaper to own than to rent, which made ownership more desirable and possible, which in turn encouraged more building, and the land on which to build was surrounding the cities."

Respondents received the lowest scores on Question 2, where only 46.2% received two or three points. To receive three points on Question 2, respondents needed to identify two reactions from factory owners to the formation of unions in the 1800s and early 1900s. The opinion was to be explained with specificity. One teacher earned a three by writing: "1) Factory owners violently broke strikes with the help of the government. 2) Factory owners hired non-union workers." A two point answer identified two reactions but was not specific enough to receive the full score. For example, "Factory workers reacted by refusing to hire employees who were unionized and denying their workers time to congregate in Union environments." Respondents who received one point either identified only one reaction or the answer was correct but not specific enough. One respondent wrote: "1) Factory owners fired or refused to hire union members. 2) [blank]." Another response was deemed too vague: "Factory owners were unhappy and fearful of labor unions. Owners saw workers as machines and wanted to control them. They were fearful of a strong labor force because they would lose control and money over workers."

Please note that many of the 1-point responses for Question 2 identified only one reaction. More respondents may have written two reactions if the lines for the answer were clearly delineated with a 1 and a 2 (Note: this has been added to the assessment).

TAH Group Results

Overall, 43 teachers completed the baseline NAEP Teacher Content Assessment in the fall of 2008. Of the 43, 18 were in the fourth/fifth grade group, 12 were in the eighth grade group, and 13 were in the 11th grade group. Each group was administered a NAEP assessment that was made especially for that grade level (the assessment for grades 4 and 5 was combined). An overview of the group results are presented in Table 8 below.

Table 8: Baseline NAEP Assessment (All Grades) Results by Mean, Median, Minimum and Maximum Score N=43

	Sum of MC Questions (10 points possible)	Sum of Short Answer Questions (30 points possible)	Total Score (40 points possible)	Total Percent Correct
Mean	9	26.1	35.1	87.8
Median	10	27.0	36.0	90.0
Minimum	4	18	22	55.0
Maximum	10	30	40	100.0

Table 8 shows that the average score on the NAEP assessments was 35.1 points out of 40 points and the median score was 36 points. The lowest score recorded was 22 points while the highest score was 40 points (100%).

The project plans to administer a follow-up assessment to all three groups of participants in the spring of 2009. The results will address GPRA Performance Measure 1.2 and will be included in the Annual Performance Report required by the United States Department of Education.