

AN EMPIRICAL ANALYSIS OF ALTERNATIVE ASSESSMENT STRATEGIES IN THE HIGH SCHOOL ECONOMICS CLASS

by Jane Lopus* and Jody Hoff**

Abstract

Three different assessment methods are used to determine the effectiveness of the curriculum package *Open and Operating: The Federal Reserve Responds to September 11* in improving the economic knowledge of 1291 California high school students. Student performance is found to vary by race and gender depending on the assessment method used: multiple choice questions, an essay question, or a group creative poster activity. Consistent with prior studies, white male students perform better on the multiple choice questions. However, these gender and racial differences disappear for the essay instrument, and females and Asians outperform others on the group creative activity. Economists who work training high school teachers are encouraged to recommend that teachers use different assessment methods to encourage the success of different types of students.

Open and Operating: The Federal Reserve Responds to September 11 is a video developed by the Federal Reserve Bank of San Francisco to inform the public about the role the central bank can play in responding to a crisis. The title refers to the Fed's press release of September 11, 2001 following the terrorist attacks: "The Federal Reserve system is open and operating. The discount window is available to meet liquidity needs." The 16-minute video describes the background and functions of the Federal Reserve System, monetary policy, how the central bank's responsibilities have evolved over time, and specifically how the Fed responded to the September 11 crisis.

After developing the video for general audiences, the Public Information Department of the San Francisco Fed designed an accompanying curriculum guide for use in high schools. The guide provides a detailed lesson plan that is divided into three main sections: an introduction with visuals, instructions for viewing the video using student handouts, and a follow-up poster activity using graphic organ-

izers. The guide relates the concepts covered to the *Voluntary National Content Standards in Economics* (NCEE 2000), and includes the transcript of the video and detailed instructions for stopping it at certain points to discuss and clarify concepts with students. The post-viewing poster assessment activity, which is completed in groups, includes a rubric for teachers to use in evaluating the poster.

This paper assesses the effectiveness of the *Open and Operating (O&O)* video and curriculum on improving economic knowledge in California high school students using three different assessment methods: multiple choice questions, an essay question, and the creative poster assessment activity. We are able to compare results from the different assessment methods across demographic categories of students to determine factors affecting achievement on the different instruments. We use data collected from 24 teachers and 1291 students, divided into control and experimental groups.¹ Students in the experimental group viewed the *O&O* video and participated in the accompanying

* Corresponding author, Professor of Economics and Director, Center for Economic Education California State University, East Bay 25800 Carlos Bee Blvd. Hayward, CA 94542 510.885.3140; fax: 510.558.8033 jane.lopus@csueastbay.edu

** District Manager Economic Education and Public Programs San Francisco Federal Reserve Bank 101 Market Street; MS 640 San Francisco, CA 94105 415.974.2952; fax: 415.974.3341 jody.hoff@sf.frb.org.

We would like to thank an anonymous referee and Marcelo Clerici-Arias and participants in an economic education research session at the annual meetings of the Allied Social Science Association, Jan. 4–6 2008, New Orleans, LA for helpful comments. We would also like to thank Peter Kennedy for econometric advice and Ben Santa Maria and Nick Haltom for data assistance.

lesson activities while students in the control group did not. All students took multiple choice pre and post tests and answered a pre and post short essay (free-response) question. Students in the experimental groups also completed the poster assessment activity as part of the *O&O* curriculum.

I. Literature Review

Prior research identifies factors affecting achievement in high school economics classes. Watts (2005) finds that taking a separate course in economics, teacher training, teacher experience, good instructional materials, and student characteristics such as cognitive skills are important predictors of student learning of economics.

Survey papers by William Walstad (1992; 2001) also address determinants of learning economics in high schools. In the *Journal of Economic Perspectives* (2001), Walstad reports that although taking a high school economics class contributes to economic understanding, high school students still manifest a “relatively low level of achievement” likely due to limited instructional time for economics compared to subjects like history that are taught over a series of years (p. 202.) Teachers who take more economics courses are more effective in teaching economics (p. 206), with about five or six courses thought to be a critical threshold level.

Walstad’s 1992 survey paper in the *Journal of Economic Literature* directly addresses the use of supplementary materials in teaching high school economics. Developing new curriculum materials and using different teaching strategies “are dominant themes in economic education work. . .” (p. 2039.) A major problem with materials used to supplement textbooks is being able to distinguish the quality materials from materials found to be unacceptable (p. 2042.) As well, using films to teach economics was generally shown to have a positive effect on achievement.

There is a fairly consistent literature showing that male students outperform female students in high school economics classes (Evans 1992, Peterson 1992, Heath 1989, Walstad and Soper 1989, Siegfried 1979.) Most of these studies use a multiple choice test such as the *Test of Economic Literacy* (Walstad and Rebeck 2001) to assess achievement. There is some evidence that the male-female differences are limited to multiple choice

tests and that female students may perform better than males on essay questions. In a study of over 3000 college and university students in the U.K., Lumsden and Scott (1987) found that male students performed better on multiple choice questions but that female students performed better on essay questions. Ferber, Birnbaum and Green (1983) found that female economics students at one U.S. university performed statistically as well as their male counterparts on essay questions, but worse on multiple choice questions. These different results on different types of questions are thought to be due to male superiority in spatial and numerical skills and female superiority in verbal skills. As well, studies that also control for race or ethnicity frequently find that white students outperform non-white students on multiple choice tests (Walstad and Soper 1989, Lopus 1990, Evans 1992.)

This paper contributes to the literature by comparing results on multiple choice questions, an essay question, and a group creative assessment activity across a large sample of high school economics students, while also exploring the effects of the *Open and Operating* curriculum package.

II. Research Design

1. Description of Sample: In an attempt to recruit an unbiased sample from the target population of California economics teachers, letters and emails describing the project and inviting teachers to participate were sent to all (approximately 1000) public high schools in California in spring 2006. Involving more than one state would introduce variation from different state standards in economics and different teacher credentialing requirements. California has a large and ethnically diverse student population, allowing us to introduce demographic controls in our study that have been found to be significant in the economic education literature. A semester course in economics has been required for high school graduation in California since 1989. Most schools offer the course to high school seniors, backed with a semester course in government. Economics is included in the *California History-Social Science Content Standards for California Public Schools*, and to “(u)nderstand the aims and tools of monetary policy and their influence on economic activity (e.g., the Federal Reserve)” is one of those standards. (California State Board of Education, 1998.)

To be eligible to participate in this study, we sought teachers who were teaching two similar economics classes during fall 2006. That is, teachers could be teaching two Advanced Placement Classes, two honors/college prep classes, two mixed ability classes, or two classes of non-college bound students. In this way one class could serve as the experimental class and the other class could serve as the control class. Having each teacher in the study use two classes eliminated the difficult problem of locating control group teachers and then having to match teachers and classes with similar characteristics.

Teachers agreed to use the *Open and Operating* video and curriculum in the experimental class in addition to the materials they normally use to teach about the Fed and monetary policy. The *O&O* video curriculum takes about two and a half hours of class time. In the control class, teachers would not use the video curriculum and would teach about the Fed and monetary policy in their usual manner, substituting other instruction on the topics so that equal amounts of time would be used to teach about the Fed and monetary policy in both classes. They agreed to administer a short pre and post test and student questionnaire in both classes. They also agreed to score the poster assignment from the experimental class² and to complete a teacher questionnaire. If all requisite materials were returned, participating teachers received a \$250 consultant fee.

Sixty-two teachers from throughout California returned the response form indicating that they would be teaching two similar economics classes in fall 2006 and that they were available to participate in the project. These 62 teachers were sent sets of pretests, student background questionnaires and instructions in early September 2006. Post-tests, two additional background questions for students, and the teacher questionnaire were sent in November 2006. Teachers were instructed to randomly assign one class to be the experimental group and one class to be the control group. Although it may have been preferable for us to randomly assign the classes to the experimental and control groups, we have no reason to believe that bias exists from the teachers assigning the classes. Forty-three teachers returned some materials and twenty-four teachers returned complete sets of usable materials for both control and experimental classes.³

The teacher questionnaire asked for information about teaching experience and background in studying economics. **Table I** reports characteristics

TABLE I.
Teacher Characteristics

Characteristic	Number of Teachers
Gender	
Male	17
Female	7
Years of teaching experience	
<10	5
11–20	10
21–30	5
> 30	4
Years of experience teaching economics	
3–5	5
7–10	7
13–20	11
> 20	1
Teaching econ is what percent of total teaching load	
10–44	4
50–75	9
80–100	11
Have advanced degree	
No	8
Yes	16
Have major, minor and/or MA/MS in economics	
No	14
Yes	10
Semester credit hours of economics	
0	2
3–10	6
12–20	7
25–30	3
>30	4
Number of non-credit economics workshops in past 5 years	
0	3
1–2	6
3–4	8
5–10	7
Attitude toward teaching economics	
Love it—one of all time favorite subjects to teach	23
It's ok, but not one of my all-time favorite subjects to teach	1
Really don't like it—one of my least favorite subjects to teach	0

n = 24 (numbers do not always total 24 due to missing data)

from the 24 teacher questionnaires. Seventeen of the teachers are male and seven are female. The teachers are generally experienced teachers and are also experienced in teaching economics. On average, teaching economics represents a major portion of their teaching assignment. Most of the teachers (16) have an advanced degree, and ten have either a major, minor, and/or advanced degree in economics. All except two of the teachers have completed some college coursework in economics, with most taking between one and four classes. As well, most of the teachers (21) have taken at least one non-credit workshop in economics in the past five years, with an average of three to four. Teachers in the sample report that they are very enthusiastic about teaching economics.

Teachers were also asked to report information about the classes in the study and the appropriateness of the *O&O* video curriculum. This information is shown in **Table II**. In describing the classes, 22 teachers taught twelfth-grade classes and the other two taught eleventh grade classes. Most of the teachers (15) taught classes of mixed ability students, although two taught Advanced

Placement classes, six taught honors or college-prep classes, and one taught non-college bound students. For the most part, teachers followed instructions calling for equal treatment in the control and experimental class. There are no differences in the time lag between teaching the unit on the Fed and monetary policy and the administration of the post test. All teachers reported randomly assigning classes to the experimental or control group or reported that they perceived no significant differences between the classes. However, two teachers reported that they spent more time teaching about the Fed and monetary policy in the experimental class despite the instructions to spend equal amounts of time in both classes. This could create a problem for the study because we want to capture the effects of the *O&O* curriculum and the different assessment methods and not more time spent teaching the content. Another potential problem is that two teachers did not follow the instructions to tell students that the post test would count toward their grade in the class, which could possibly result in students not taking the assessment seriously.

TABLE II.
Information about Classes in the Study

Course Characteristic	Number of Teachers / (or average)
Grade level of classes in study	11th: 2 12th: 22
Level of classes in the study	Non-college bound: 1 Mixed ability: 15 College Prep/Honors: 6 Advanced Placement: 2
Randomly assigned classes to experimental or control class, or perceive no differences between them	Yes: 24 No: 0
Spent more time teaching about the Fed and monetary policy in the experimental class	Yes: 2 No: 22
Students were told that post test would count toward their grade	Yes: 22 No: 2
Average number of days between teaching the content and giving the post test	Experimental: 2.74 Control: 2.78
Appropriateness of level of <i>O&O</i> video for use with classes	too difficult: 1 about right: 23 too easy: 0
Appropriateness of level of <i>O&O</i> curriculum package video for use with classes	too difficult: 0 about right: 23 too easy: 1

n = 24

Because the *O&O* video was designed for adult audiences, it is interesting that most teachers (23) reported that the video was an appropriate level for their classes. Only one reported that it was too difficult and one reported that the accompanying curriculum was too easy.⁴

Table III reports student characteristics from the questionnaire distributed with the pretest and two additional questions sent out with the post test. Students in the sample are about evenly divided between male (51 percent) and female (49 percent). They are ethnically diverse, with 42 percent being white, 22 percent Latino/Hispanic/Chicano, 15 percent Asian/Pacific Islander, four percent Black/African American, and 17 percent other or mixed race. Most (80 percent) report that English is their preferred language, although four percent prefer another language and 17 percent are bilingual. Most (85 percent) plan to attend a two or four-year college or university after graduating from high school. They expect to earn high grades in their economics classes, but are not all that excited about taking the class, with 68 percent indicating that they think it will be ok, but not one of their favorite subjects. Most (86 percent) have not taken a prior course in economics or are not aware that they have studied it in other classes. Students self-report an average grade point average of 3.0. Although self-reported grades are likely to be inflated with weaker students more likely to inflate their grades (Maxwell and Lopus 1994), school-provided GPA data were not available for this study. The highest education level attained by a parent is widely dispersed, with 11 percent not graduating from high school, 12 percent earning a high school degree, 30 percent attending some college or earning a two-year degree, 30 percent earning a four-year degree, and 17 percent earning an advanced degree.⁵ Six-hundred- fifty-two students were in classes designated as experimental and 632 were in control classes.

2. Pre and Post Test Instruments: Because the video and curriculum address economic concepts relating almost exclusively to the Fed and monetary policy, we searched for an existing multiple choice test instrument focusing on these concepts to use as pre and post tests to test the effectiveness of the *Open and Operating* program for high school students. It would have been desirable to have an appropriate valid, normed and reliable instrument to use such as the *Test of Economic*

Literacy (TEL) (Walstad and Rebeck 2001.) Although we identified several appropriate questions from existing instruments, we were unable to locate an appropriate test. Therefore, in the absence of such a test we sought to develop an instrument with strong content validity. (Becker and Walstad 1987 p. 87–91.)

Using both the *Voluntary National Content Standards in Economics* and the California State Standards, we developed a list of concepts related to money and banking that would be covered in a high school economics class. We then located six existing questions on the *TEL* that addressed these concepts. By using these questions in our instrument, we are able, to a degree, to ground our test with the reliability and norming features of the *TEL*. We also located existing questions from the Kids Quiz on the Federal Reserve website (www.federalreserve.gov/kids/default.htm) and from released *Advanced Placement* exams. Because this collection of questions did not fully address the identified topics in the state and national standards, we added selected questions from text book test banks.

We then formed an advisory group of three teachers to preview the sample questions. These teachers were not participants in the study and all have experience teaching both AP and regular economics classes. Based on their comments, we ended up with battery of 20 multiple choice questions covering a range of concepts related to the Fed and monetary policy and one short essay question for the pretest. The essay question was:

“Imagine that there is a crisis in the economy. (The crisis could be high unemployment, high inflation, a banking panic, or something else that you choose to write about.) Write a short article (one or two paragraphs) for your local newspaper describing how the Federal Reserve System could respond to this crisis to help the economy.”

The post test consisted of the same multiple choice questions (but in a different order and with the order of the responses also changed) and the same essay question.

Although students answered 20 multiple choice questions, one of the questions taken from the *TEL* was not scored due to ambiguity in the response⁶ so the final multiple choice battery consisted of 19 questions. Students entered answers to multiple choice questions on Scantron forms that were machine scored. An advisory group of teachers was

TABLE III.
Student Characteristics

Characteristic	Percent of Sample
Gender:	
Male	51.0
Female	49.0
Ethnicity	
White	42.1
Latin American / Hispanic / Chicano	22.1
Asian / Pacific Islander	15.2
Black / African American	3.9
Other or Mixed	16.7
Preferred Language	
English	79.5
Other	3.7
Equally proficient in English and another language	16.8
Plans after High School	
Full time job	1.7
Military	3.8
Vocational / technical school	4.3
Two year college	28.6
Four year college	56.1
Other	5.6
Expected Grade in Economics Class	
A	49.3
B	37.0
C	11.6
D or F	.8
Attitude Toward Studying Economics	
Very excited and think it will be one of my favorite subjects	25.6
Think it will be ok but not one of my favorite subjects	67.8
Don't like it and think it will be one of my least favorite subjects	6.5
Prior Economics Studied	
Never studied economics before	85.8
Have taken a prior course in economics	2.5
Have studied economics in other classes	11.7
Self-reported High School Grade Point Average	
> 4.0 (due to honors and AP credit)	37
3.6–4.0	205
3.0–3.5	443
2.0–2.9	335
1.1–1.9:	27
Highest Education Level Attained by a Parent	
Not a high school graduate	11.3
High school graduate	11.8
Some college	20.1
Graduated from a two-year college	10.0
Graduated from a four-year college	30.1
Had an advanced degree	16.6

$n = 1047-1290$

hired to score responses to the essay question. Scores assigned were 0 (no response or totally irrelevant response), 1 (confused, incorrect information, but some indication of knowledge about the Fed's role), 2 (mostly correct information; appears to have basic knowledge about the role of the Fed) and 3 (correctly describes the role of the Fed in addressing the crisis identified, including details.)

Because the *Open and Operating* curriculum contains a poster assessment activity, we have an opportunity to investigate a measure of criterion-related validity, which can be determined by connecting performance on one type of (economics) test with a related measure (Becker and Walstad 1987 pp. 84.) If it can be argued that the poster activity in the *O&O* curriculum is a valid and reliable assessment technique, its scores can be compared to the post test scores on the multiple choice and essay questions. The poster assignment consisted of groups of two or more students depicting a visual summary of the *O&O* video. Teachers were given a rubric with which to score the posters in terms of presentation, content, and creativity. Scores could range from 0–12.

III. Teaching Styles and Learning Styles

The questionnaires also asked teachers what methods and materials they perceived to be beneficial to their teaching, and asked students what methods and materials they perceived to be benefi-

cial to their learning. A similar list of methods and materials was provided on the teacher and student questionnaires. It is interesting to note the differences reported by teachers and their students. This comparison is shown in **Table IV**. In each case teachers were significantly more likely to identify the method or materials as being helpful to their teaching than students were to say it was helpful to their learning. However, both teachers and students most often identified videos as being helpful and least often identified computer software as being helpful. If rankings may be inferred from the teacher and student responses, it is interesting that students rank “creative activities” third whereas teachers rank them eighth, and teachers rank lectures tied for their second favorite whereas students rank lectures third from the bottom. The implications of these differences in teaching and learning styles are a topic for future research.

IV. Testing Results

Table V compares pre and post test scores and change scores (differences between pre and post test scores) for experimental and control classes on the multiple choice and essay questions. For the multiple choice questions, the pretest scores of the experimental and control groups were not significantly different from each other. However, the post test scores, and therefore the change score, were different for the experimental and control groups.

TABLE IV.
Teaching and Learning Styles

Method / Materials (0 = No; 1 = Yes)	Teachers: Helpful in Teaching mean/(std. dev.)	Students: Helpful in learning mean/(std. dev.)	t (sig)
Videos	.95 (.22)	.69 (.46)	−3.48*** (.001)
Computer Software	.59 (.50)	.30 (.46)	−3.95*** (.000)
The Internet	.82 (.39)	.48 (.50)	−4.19*** (.000)
Newspaper, Magazines	.87 (.34)	.31 (.46)	−7.47*** (.000)
Lectures (including PowerPoint)	.92 (.27)	.41 (.49)	−6.46*** (.000)
Discussions	.90 (.31)	.64 (.48)	−3.36*** (.001)
Games, Simulations	.92 (.27)	.59 (.49)	−4.24*** (.000)
Creative Activities	.79 (.41)	.63 (.48)	−2.09** (.040)
Group Cooperative Learning Activities	.82 (.39)	.53 (.50)	−3.74*** (.000)
<i>n</i>	24	1290	

** significant at .05 level

*** significant at .01 level

TABLE V.
Pre and Post Test Results

	Experimental mean/(stnd dev)	Control mean/(stnd dev)	t (sig)
Multiple choice pretest score (out of 19) <i>n</i> = 630 – 633	7.97 (2.62)	8.00 (2.49)	–.187 (.852)
Multiple choice post test score (out of 19) <i>n</i> = 569 – 587	13.08 (3.30)	12.49 (3.56)	2.95*** (.003)
Multiple choice change score <i>n</i> = 561 – 572	5.07 (3.62)	4.51 (3.51)	2.63*** (.009)
Essay pretest score (out of 3) <i>n</i> = 620 – 640	.170 (.447)	.240 (.554)	–.248 (.013)
Essay post test score (out of 3) <i>n</i> = 534 – 586	1.06 (.976)	1.02 (.941)	.738 (.461)
Essay change score <i>n</i> = 520 – 574	.887 (.983)	.810 (.927)	1.332 (.183)
Poster assessment score <i>n</i> = 574	8.86 (2.24)	—	—

*** significant at .01 level

Students in the experimental classes scored significantly higher on the multiple choice post test than students in the control classes who did not participate in the *O&O* curriculum.⁷ This implies that the *O&O* curriculum is superior to the other materials that teachers used to teach about the Fed and monetary policy in their control classes, as measured by multiple choice questions. However, it should be noted that the teachers self-selected into the study and were overwhelmingly enthusiastic about teaching economics. This may have led to enthusiasm for the new materials, affecting student performance in some classes.

Table V reveals that there are no overall differences between the experimental group and control group for the essay question on the pretest, post test, or change score. That the *O&O* curriculum affected the multiple choice scores but not the essay scores is perhaps surprising, because viewing the video may be expected to provide a context for the content that could translate into a more reasoned discussion on the essay question. That this did not occur may suggest that the video-based curriculum allows students to identify correct answers more readily, but not to conduct a better independent analysis when given an open-ended question. Table V also reports the average poster score for the students in the experimental classes. Control group scores are not available for the poster assessment activity since the poster

activity was part of the *O&O* curriculum and therefore was not completed by the control group.

Because teachers used their own classes as controls, it is interesting to look at the multiple choice and essay scores by teacher broken down into experimental and control classes. This information is provided in **Table VI** and **Table VII**. The first columns assign a teacher number and indicate the level of the course: Advanced Placement (AP), college prep or honors (CP-H), mixed ability (Mixed) or non-college bound (NonCol). The next columns give pretest, post test, and change scores by teacher. The last column indicates if the experimental class performed significantly better than the control class, determined by t-tests comparing average gains (change scores) between the pretest and post test in the different classes for each teacher. First looking at Table VI for the multiple choice questions, we see that all of the classes had higher average post test scores than pretest scores and for seven of the 24 teachers, the students in the experimental classes who took part in the *O&O* curriculum had significantly higher change scores than those in the control classes. These were college prep/honors classes (two) or classes with mixed ability students (five.) None of the teachers had control classes with significantly higher change scores than the experimental classes. This reinforces the conclusion that the *O&O* curriculum significantly improved students'

TABLE VI.
Multiple Choice Scores by Teacher

Teacher / Class Level	Control Class Average (19 possible)			Experimental Class Average (19 possible)			t-test (Experimental -Control Change Scores)
	Pretest	Post test	Change	Pretest	Post test	Change	
1 AP	10.00	16.85	6.85	11.00	17.00	6.00	-1.242
2 Mixed	7.69	10.96	3.27	8.62	11.42	2.81	-.571
3 NonCol	7.17	12.08	4.92	6.81	13.06	6.25	1.06
4 Mixed	9.07	11.75	2.68	8.73	11.62	2.88	.247
5 Mixed	8.07	10.27	2.20	8.44	12.48	4.04	1.814*
6 CP-H	7.90	10.57	2.67	8.24	14.00	5.76	3.985***
7 CP-H	8.28	10.44	2.16	7.46	11.23	3.77	2.187**
8 Mixed	7.87	11.83	3.97	8.11	13.63	5.52	1.881*
9 CP-H	7.93	12.83	4.90	8.16	13.88	5.72	1.210
10 Mixed	6.50	7.43	.93	6.68	7.33	.67	-.186
11 Mixed	7.19	14.89	7.70	6.54	13.69	7.15	-.649
12 AP	8.00	14.76	6.76	7.93	12.87	4.93	-1.587
13 Mixed	9.36	13.50	4.14	9.27	14.58	5.31	1.547
14 CP-H	8.50	13.42	4.92	9.35	13.66	4.31	-.542
15 Mixed	7.48	14.32	6.84	5.91	15.70	9.78	3.624***
16 Mixed	7.66	13.03	5.37	6.34	13.69	7.34	2.655***
17 Mixed	7.58	13.79	6.21	7.23	12.71	5.48	-.908
18 CP-H	8.64	16.86	8.23	7.96	15.89	7.93	-.295
19 Mixed	9.32	13.23	3.91	8.91	12.57	3.65	-.286
20 Mixed	7.17	8.48	1.30	7.83	9.87	2.04	.351
21 Mixed	9.00	12.07	3.07	9.74	14.57	4.83	1.944*
22 Mixed	6.65	8.95	2.30	6.96	10.00	3.04	.703
23 Mixed	6.80	12.00	5.20	7.37	12.32	4.95	-.297
24 CP-H	9.00	14.56	5.56	8.18	12.82	4.66	-1.121

* significant at .10 level

** significant at .05 level

*** significant at .01 level

multiple choice scores, both in the aggregate and when viewed by teacher.

Table VII, which reports essay scores by teacher, tells a somewhat different story about the effects of the *O&O* curriculum. Here in all but one class students again scored higher on average on the post test than on the pretest, but for four of the teachers, the experimental classes had significantly larger change scores than the control classes and for five of the teachers, the control classes had significantly larger change scores than the experimental class. These differences cancel out in the aggregate, explaining why no overall differences in essay scores are found in Table V. These differential teacher effects may indicate that the video-based

curriculum may interact better with some teaching styles than with others, leading some classes to exhibit better explanatory power on the essay question. Further reinforcing this idea, it is interesting that only one of the teachers (# 21) whose students showed significant differences in the essay scores also showed significant differences in the multiple choice scores, and it was the control class that had the higher essay change scores.

V. Regression Analysis

To determine the effects of the *O&O* video and curriculum on learning in the high school classes

TABLE VII.
Essay Scores by Teacher

Teacher / Class level	Control Class Average (Scores range from 0–3)			Experimental Class Average (Scores range from 0–3)			t-test (Experimental – Control Change Scores)
	Pretest	Post test	Change	Pretest	Post test	Change	
1 AP	.714	2.214	1.500	1.000	2.278	1.278	–.594
2 Mixed	.087	1.130	1.043	.250	1.143	.893	–.555
3 Non Col	.167	.333	.166	.118	.471	.353	.755
4 Mixed	.917	1.333	.416	.190	1.524	1.334	2.434**
5 Mixed	.000	.286	.286	.185	.630	.445	.526
6 CP-H	.267	1.067	.800	.138	1.897	1.759	3.145***
7 CP-H	.333	1.458	1.125	.154	1.192	1.038	–.345
8 Mixed	.103	.828	.725	.037	.889	.852	.699
9 CP-H	.688	1.594	.906	.235	2.118	1.883	3.895***
10 Mixed	.143	.000	–.143	.000	.333	.333	1.561
11 Mixed	.000	.741	.741	.000	.731	.731	–.069
12 AP	.150	1.300	1.150	.267	1.800	1.533	1.135
13 Mixed	.364	1.636	1.272	.385	1.500	1.115	–.605
14 CP-H	.067	1.067	1.000	.174	1.000	.826	.651
15 Mixed	.083	1.500	1.417	.000	1.708	1.708	.981
16 Mixed	.031	.781	.750	.143	.486	.343	–2.240**
17 Mixed	.394	.939	.545	.161	1.194	1.033	2.186**
18 CP-H	.000	.773	.773	.000	.308	.308	–2.774***
19 Mixed	.286	.952	.666	.333	.542	.209	–1.618
20 Mixed	.125	.500	.375	.083	.375	.292	–.446
21 Mixed	.071	.714	.643	.217	1.522	1.305	2.178**
22 Mixed	.050	.150	.100	.087	.174	.087	–.099
23 Mixed	.200	1.000	.800	.105	.368	.263	–2.886***
24 CP-H	.200	1.640	1.440	.045	.773	.728	–3.062***

* significant at .10 level

** significant at .05 level

*** significant at .01 level

while controlling for other variables, we estimated a series of regression equations investigating factors affecting the different types of assessment outcomes. Following the economics education production function literature, we control for teacher and student characteristics found to be significant in other studies involving learning economics in high schools. **Table VIII** reports regression results for factors affecting the post test multiple choice score (Column 2), the post test essay score (Column 3), and the score on the poster activity (Column 4). The multiple choice and poster equations were estimated using ordinary least squares (linear) regression since the scores range from 1–19 and 0–12 respectively, while the essay equation was esti-

mated using ordered logistic regression since the scores range from 0–3.

First investigating the multiple choice test results, the experimental class scored significantly higher than the control class even when controlling for the fact that some teachers spent additional time on the content in the experimental class and other factors. This result is robust across different specifications. The level of the class is also significant, with higher level classes outperforming those with mixed ability and non-college bound students. Student gender is significant, with males scoring higher than females, consistent with prior studies. White students, the omitted category in the ethnicity category, scored significantly

TABLE VIII.
Regression Results

Independent Variables	Dependent Variable: Post test multiple choice score	Dependent Variable: Post test essay score	Dependent Variable: Poster activity score
	B Std. Error (sig)	B Std. Error (sig)	B Std. Error (sig)
Score on multiple choice pretest	.239*** .038 (.000)	----	----
Score on essay pretest	----	.732*** .131 (.000)	----
Experimental Class (= 1; control = 0)	.389** .183 (.034)	.087 .122 (.479)	----
Years of teaching economics	.001 .012 (.909)	.037*** .008 (.000)	-.014 .012 (.266)
Level of Class (low to high)	.282*** .096 (.003)	.364*** .064 (.000)	.234** .103 (.023)
Student gender (0 = female, 1 = male)	.430** .185 (.020)	.179 .124 (.148)	-.685*** .194 (.000)
Asian	-.229 .278 (.410)	-.085 .184 (.645)	.680** .297 (.022)
Black	-1.108** .472 (.019)	-.583 .327 (.074)*	.294 .528 (.577)
Hispanic	-1.546*** .278 (.000)	-.244 .186 (.189)	-.169 .294 (.565)
Mixed race/other	-.312 .262 (.233)	.086 .173 (.617)	-.091 .271 (.736)
High school GPA	1.582*** .177 (.000)	1.030*** .123 (.000)	.770*** .179 (.000)
Parents' educational level	.133** .064 (.037)	.105** .042 (.013)	.011 .067 (.867)
Student attitude toward economics (1 = like, 3 = don't)	-.740*** .173 (.000)	-.140*** .051 (.006)	-.411** .180 (.023)
More time in experimental class	.704** .354 (.047)	.962*** .233 (.000)	-.066 .365 (.857)
Constant	6.332*** .742 (.000)	ScorPostEs=0: 3.684*** ScorPstEs=1: 5.520*** ScorpostEs=2: 7.181***	7.203*** .709 (.000)
	<i>n</i> = 982 <i>r</i> ² = .288	<i>n</i> = 963 <i>r</i> ² = .249	<i>n</i> = 485 <i>r</i> ² = .143

* significant at .10 level
 ** significant at .05 level
 *** significant at .01 level

higher than Blacks and Hispanics. The student's self-reported high school GPA and the educational level attained by the parents are both significantly correlated to performance on the multiple choice

test, as is the student's having a more positive attitude toward the subject of economics.⁸

Column three in Table VIII reports ordered logistic regression estimates with the essay post

TABLE IX.
Correlations between Different Test Instruments

Test Instruments	Pearson Correlation	Sig.	N
Poster and Essay Post Test	.149	.000	541
Multiple Choice Post Test and Poster	.120	.003	544
Multiple Choice Post Test and Essay Post Test	.395	.000	1110

test score as the dependent variable and the same independent variables as in the multiple choice regression, except that the essay pretest score was used instead of the multiple choice pretest score. Results similar to the multiple choice findings exist for the pretest score, the class level, high school GPA, parents' educational level, student attitude toward economics and whether more time was spent teaching about the subject matter. Interesting differences also exist when comparing the multiple choice results to the essay results. As expected from Table V, there are no significant differences between the experimental and control classes on the essay scores. However, students whose teachers are more experienced teaching economics score higher than students whose teachers have less experience. Consistent with implications from other studies of college students, males do not perform significantly better than females on the essay question (Lumsden and Scott 1987, Ferber, Birnbaum, and Green 1983.) The performance between Hispanic students and white students is not significant, and that between Blacks and whites only marginally significant.

We next investigate the scores on the creative group poster activity shown in Column 4 of Table VIII and compare these findings to those of the multiple choice and essay questions. Consistent with the multiple choice and essay results, students in more advanced classes outperform those in e.g. non-college bound classes. High school GPA and student attitudes affect scores. However, on the creative poster activity the ethnicity and gender results are very different from those on the multiple choice part of the test: Asian students outperform whites and females score higher than males. This may indicate that certain students are more likely to excel on an creative type of activity, or it may also be a reflection of some students working better in groups, since the multiple choice and essay assessments were completed individually

where the poster activity was a group project and received a group grade.

VI. Correlations between Different Test Instruments

Table IX reports correlation coefficients between the poster and essay assessments, the multiple choice and poster assessments, and the multiple choice and essay assessments. In each case the scores were positively and significantly correlated to each other at the .01 confidence level. This suggests that using creative assessment activities such as the poster assignment in the *O&O* curriculum provides similar information to the teacher as do more traditional forms of assessment such as multiple choice tests or essay tests. Since "creative activities" are among students' favorites, as shown in Table IV, using more creative assessment activities may serve to improve both student attitudes toward economics and therefore student achievement in economics. As well, regression results indicate that different types of student learners score higher on different types of assessment activities. White and male students score higher than others on multiple choice tests, but no gender differences are observed on the essay question and ethnicity differences almost disappear. Different gender and ethnicity effects are observed for the creative poster activity, with female and Asian students outperforming others. Therefore, if teachers assess learning using different methods, different groups of students may have the opportunity to demonstrate knowledge gained and therefore to succeed in the economics class.

VII. Summary and Conclusions

Using three different assessment methods, this paper provides findings on the effects of the video and curriculum *Open & Operating: The Fed*

Responds to September 11 on learning in the high school economics classes of 24 teachers and 48 classes of students in California. Teachers in the sample are experienced economics teachers who enjoy teaching economics. Students in the sample are primarily high school seniors and are ethnically and socioeconomically diverse. Most plan to attend college and expect high grades in their economics course, although they don't expect to particularly like it. Students are significantly less enthusiastic than their teachers about methods and materials that may help their learning, perhaps indicating a disconnect between teaching styles and learning styles.

Pre and post test results indicate that students who participated in the *Open & Operating* curriculum score significantly higher on a 19-question multiple choice test than students who learned material about the Fed and monetary policy in more traditional ways, controlling for factors related to teaching and learning economics. The *O&O* curriculum was not found to affect overall scores on an essay question, with four experimental classes having larger change scores and five control classes having larger change scores. The *O&O* curriculum also contains a poster assessment activity, allowing comparison of the three different assessment methods.

Although there are similarities, regression results indicate that different factors are correlated with higher scores and more learning on the different types of assessment activities. Although no gender differences are found with respect to the essay question, males score higher than females on multiple choice questions and females outperform males on the creative poster assessment. There are ethnic/racial differences on the different types of assessment activities, with whites outperforming Blacks and Hispanics on the multiple choice questions and Asians outperforming others on the poster activity.

The findings suggest that using a video curriculum such as the Fed *O&O* program to teach economics instead of more traditional methods can help to improve student achievement in economics with respect to multiple choice questions. The findings also suggest that economists who work in training high school economics teachers should recommend that they use different types of assessment activities to promote success for different types of student learners.

NOTES

1. It should be noted that since this study involves human subjects, IRB approval for conducting the study was received through the California State University East Bay Institutional Review Board. The study was found to be exempt from IRB Review.
2. Note that only students in the experimental classes completed the poster activity because this activity was part of the *O&O* curriculum, which control group students did not experience.
3. Many of the original 62 teachers who responded dropped out because their teaching schedules changed over the summer, which is not uncommon in high schools. Of the 43 teachers who returned materials, some had problems such as our not being able to match student pre and post information, their using different classes for pre and post testing, their not completing the poster assignment, or otherwise not following instructions. Thus our empirical sample consists of information from 24 teachers and 48 classes.
4. However, several teachers who checked the box saying the video was an appropriate level for their classes added written comments saying that it was hard to understand or that it moved too quickly.
5. The questions on GPA and parent education have fewer responses (1047-1088) compared to the other questions (1286 - 1290.) These questions were sent out with the post test and were supposed to have been printed with the post-test essay question. However, due to a printing error, some students received a page with a blank box that should have contained the questions. Teachers were then sent the extra questions separately and asked to return them, but there were data lost in the process. Also, there is some chance that post-test change scores may be lower due to the printing error, as some students wrote their essay responses in the blank box, which was smaller than the space provided on the pre-test. We do not believe that the data loss or the possibility of lower post test essay scores occurred in a systematic way that would affect the results.
6. The question asked "Which makes up the major portion of the money supply in the US?" but did not distinguish between M1 and M2. Possible responses were gold, currency and

coins, deposits in checking accounts, and deposits in savings accounts. Currently, currency makes up the major portion of M1 (52 percent) but deposits in savings accounts make up the largest part of M2.

7. Since the experimental group scored about half a question better on the multiple choice post test than the control group, it may be interesting to investigate if there was a particular question that caused this difference. This information is not available at this time.
8. A variable controlling for whether the post test counted towards the students' grades is not included in the model because it was consistently found to have the wrong sign for the two classes in the sample who were not told that the post test would count toward their grade. That is, these two classes had large and positive change scores indicating that this did not affect their motivation to do well on the post test.

References

- Becker, William E. Jr. and William B. Walstad, ed. *Econometric Modeling in Economic Education Research*. Kluwer-Nijhoff Boston MA, 1987.
- History-Social Science Content Standards for California Public Schools*. California State Board of Education, October 1998.
- Evans, Mark O. 1992. An Estimate of Race and Gender Role-Model Effects in Teaching High School. *The Journal of Economic Education*, Vol 23, No. 3 (Summer) pp. 209–231.
- Ferber, Marianne A., Bonnie G. Birnbaum and Carole A. Green. Gender Differences in Economic Knowledge: a Reevaluation. *The Journal of Economic Education*, Vol. 14, No. 2 (Spring 1983) pp. 24–37.
- Heath, Julia A. An Econometric Model of the Role of Gender in Economic Education. *The American Economic Review Papers and Proceedings* Vol. 79, No 2 (May, 1989), pp. 226–230.
- Lopus, Jane S. Do Additional Expenditures Increase Achievement in the High School Economics Class? *The Journal of Economic Education*. Vol. 21, No. 3 (Summer 1990) pp. 277–286.
- Lumsden, Keith G. and Alex Scott. The Economics Student Reexamined: Male-Female Differences in Comprehension. *The Journal of Economic Education* Vol. 18, No. 4 (Fall 1987) pp. 365–375.
- Maxwell, Nan L. and Jane S. Lopus. The Lake Wobegon Effect in Student Self-Reported Data. *The American Economic Review Papers and Proceedings*, Vol. 84 No. 2 (May 1994) pp. 201–205.
- Peterson, Norris A. The High School Economics Course and Its Impact on Economic Knowledge. *The Journal of Economic Education*. Vol. 23, No. 1 (Winter 1992) pp. 5–16.
- Saunders, Phillip and June Gilliard, eds. 1995. *A Framework for Teaching Basic Economic Concepts: With Scope and Sequence Guidelines. K-12*. New York: Joint Council on Economic Education.
- Siegfried, John J. Male-Female Differences in Economic Education: A survey. *The Journal of Economic Education*, Vol. 10, No. 2 (Spring 1979), pp. 1–11.
- Survey of the States: Economic and Personal Finance Education in Our Nation's Schools in 2004*. National Council on Economic Education, New York NY, 2005.
- Voluntary National Content Standards in Economics*. National Council on Economic Education, New York NY, 2000.
- Walstad, William B. and Ken Rebeck. *Test of Economic Literacy Third Edition Examiner's Manual*. National Council on Economic Education, New York NY, 2001.
- Walstad, William B. Economic Education in U.S. High Schools. *The Journal of Economic Perspectives*, Vol 15, No. 3 (Summer, 2001) pp. 195–210.
- Walstad, William B. Economics Instruction in High Schools. *Journal of Economic Literature*, Vol. 30, No. 4 (Dec. 1992), pp. 2019–2051.
- Walstad, William B. and John C. Soper. What is High School Economics? Factors Contributing to Student Achievement and Attitudes. *The Journal of Economic Education*. Vol. 20, No. 1 (Winter 1989) pp. 23–38.
- Watts, Michael. *WHAT WORKS: A Review of Research on Outcomes and Effective Program Delivery in Precollege Economic Education*. National Council on Economic Education, New York, NY, 2005.