

Instructional Research Group

**Testing the Effectiveness of Professional Development in Academic
Vocabulary on Observed Teaching Practice and Academic
Vocabulary Knowledge of English Learners in 8th Grade Social
Studies Classrooms**

**Effective Teachers and Effective Teaching – Goal 3 Replication
Study
(R305A150463)**

Final Performance Report

I. Accomplishments

A. What are the major goals of the project?

Using a research design that meets the IES What Works Clearinghouse evidence standards, researchers will report the results of analyses examining the impact of the *Teacher Study Group (TSG)* professional development (PD) program targeting eighth grade American History teachers working with a large number of current or former English learners. Researchers will report statistical significance and effect sizes of the PD impact on teacher outcomes (i.e., observed vocabulary instruction) and on student outcomes (i.e., academic vocabulary and content knowledge).

Research questions include:

1. To what extent does teacher participation in *TSG*, compared to the “business as usual” condition, result in greater use of academic vocabulary instruction aligned with the *TSG* PD model after the completion of the program?
2. To what extent does teacher participation in *TSG*, compared to the “business as usual” condition, result in better student academic vocabulary?
3. To what extent does teacher participation in *TSG*, compared to the “business as usual” condition, result in better student content knowledge? (exploratory)
4. To what extent does teacher participation in *TSG*, compared to the “business as usual” condition, result in better student grades in their American History class? (exploratory)

B. What was accomplished under these goals?

The *TSG* PD program resulted in significant impacts at the teacher level on observed teaching practice as measured by *Observation Measure for Vocabulary Instruction (OMVI)*; $g = .809$; $p = .0001$). Specifically, the Hedges’ g for two *OMVI* subscales was $.926$ ($p < .0001$) on the *Teacher-Directed Vocabulary Instruction* and $.650$ ($p = .0021$) on the *Interactive Vocabulary Instruction Scale*.

There were, however, no impacts at the student level for the main student sample or for the ELL subsample on the *Academic Vocabulary* measure or the *Assessment of Social Studies Knowledge (ASK) Content Knowledge* measure. Effect sizes ranged from $.02$ to $.07$ and were non-significant. We were unable to explore the impact of the *TSG* PD program on student grades as most of the schools in the study did not provide these data.

On the positive side, we heard many positive comments about the benefits of the *TSG* PD program from some participants. For example, in one school district in Texas, teachers indicated that their students benefitted from the program and that this was evident from their performance on vocabulary tests routinely given in American History classes.

Significant impacts at the teacher level demonstrate that the *TSG* PD program can successfully bring about change in the vocabulary teaching practices of 8th grade social studies teachers. Teachers were able to take what they learn in the PD sessions and apply it in their classrooms.

We hypothesize that the lack of impacts at the student level could be a result of teachers not consistently using the practices they learned to teach vocabulary to their students. Informal

interviews with teachers indicated that they were overwhelmed with the level of effort that was required to develop the materials to supplement texts and other readings (e.g., developing student-friendly definitions, identifying examples and non-examples, creating activities to reinforce student learning) that were necessary to implement the rich vocabulary instructional practices they learned in the PD sessions. Another potential reason was evident from the informal interviews. Teachers indicated that they were under pressure from competing demands and had to balance the extent to which they devoted time for teaching vocabulary against teaching all the history content required for state testing.

We think the next step in this line of research would be to determine if teachers would teach more words using the effective instructional practices they learned during the PD sessions if they were also given curricular material that can easily be used with their students and current texts. Professional development studies appear to be more successful when tools or materials are provided to teachers (Schwanenflugel, et al., 2010).

Activities Completed Since the Last APR

Based on the last APR, the goals for the current reporting period were to:

1. Collect student demographic data from the school districts;
2. Double check and clean student and teacher data entry;
3. Conduct analysis of the teacher observation data using the *Observation Measure for Vocabulary Instruction (OMVI)*; Gersten et al., 2010);
4. Conduct impact analysis;
5. Prepare papers and presentations to disseminate results; and
6. Submit the Final Report for Reporting Period 5 (September 2020).

We completed these activities with the exception of preparing papers and presentations to disseminate results. It took us an extraordinarily long time to gather student demographic data from districts. Despite our efforts, we were unsuccessful in collecting data for all the students. The data we received was often incorrect, and we had to follow up with schools to get the correct data. Consequently, these setbacks along with fallout from COVID-19 delayed the impact analysis and dissemination efforts. We are in the process of drafting a paper for publication and plan to disseminate the findings at future research conferences.

In the next section, we describe the activities accomplished since the beginning of the project (including teacher- and student-level impact analyses).

Activities Completed Since the Beginning of the Project

Sample

Across Cohorts 1 (2016–2017) and 2 (2017–2018), 61 middle schools (31 T, 30 C) from 17 districts in 6 states (Arizona, California, Colorado, Florida, Georgia, and Texas) participated in the study. A total of 145 eighth-grade social studies teachers (76 T, 69 C) and their 5,251 students (2,636 T, 2,615 C) from these 61 middle schools agreed to participate in the study.

Schools were randomly assigned, blocked by district, to treatment and control conditions using an Excel Visual Basic algorithm.

A total of 7 schools attrited from the study (6 T, 1 C). The main reasons for school-level attrition included (a) district leaders changing their mind about participating in the study and staffing; and (b) scheduling changes making it difficult for school personnel to participate in the

PD. The post-attrition sample includes 54 schools (25 T, 29 C). Overall attrition at the school level is 11.5%; differential attrition is 16.0%. This would be considered high attrition based on What Works Clearinghouse Standards (What Works Clearinghouse [WWC], 2020).

There were 145 teacher participants at the time of randomization across Cohorts 1 and 2 (76 T, 69 C). An additional 4 teachers joined the study after randomization (all from control schools). A total of 26 teachers attrited from the study, leaving a final sample of 123 teachers (59 T, 64 C). Of the 26 teachers, 13 teachers attrited as their district/school dropped out of the study. The remaining left for a variety of personal reasons. Overall attrition at the teacher level is 15.2%; differential attrition is 15.1%. This would be considered high attrition based on What Works Clearinghouse Standards (What Works Clearinghouse [WWC], 2020).

We randomly selected 2 classes per teacher for districts using passive consent and 3 classes per teacher for districts using active consent. This resulted in a total of 5,251 students (2,636 T, 2,615 C) across Cohorts 1 and 2. A total of 4,137 students (1,929 T, 2,208 C) had complete pretest and post-test data. The final analytic sample used for impact analysis excludes honors students and students with IEPs/504 from both conditions, resulting in a total of 4,227 students (2,162 T, 2,065 C) across Cohorts 1 and 2.

Baseline Equivalence

At the school level, there were no statistically significant differences between the treatment and control group schools in the analytic sample at baseline on any of the demographic variables. See Table 1.

Table 1. School Demographic Characteristics (2016–2018 School Years)

	Total Analytic Sample (<i>N</i> = 54)			
	<i>TSG</i> (<i>n</i> = 25)	Control (<i>n</i> = 29)	<i>t</i> (<i>df</i>)	<i>p</i>
	Mean % (<i>SD</i>)	Mean % (<i>SD</i>)		
Students reading at Proficient level or above	62.86 (15.01)	60.62 (20.31)	.46 (52)	.650
English language learners	23.56 (13.32)	22.96 (11.42)	.18 (52)	.857
Economically disadvantaged ^a	74.60 (16.67)	77.58 (15.25)	-.68 (51)	.501
Race/ethnicity				
African American/Black	11.45 (12.01)	15.58 (21.55)	-.89 (46.28) ^b	.381
American Indian/Alaska Native	0.62 (1.55)	0.48 (0.90)	.39 (38.28) ^b	.700
Asian	2.64 (4.88)	2.07 (3.84)	.49 (52)	.628
Hispanic/Latino	66.92 (24.77)	68.79 (24.49)	-.28 (52)	.782
Pacific Islander	0.08 (0.13)	0.04 (0.09)	1.16 (42.43) ^b	.255
Multiracial	1.20 (1.25)	1.20 (1.04)	.00 (52)	1.000
White	16.82 (19.83)	11.48 (11.29)	1.19 (37.88) ^b	.241

^aAnalytic sample = 53 schools (25 T and 28 C) because 2016–2017 data is unavailable for 1 school from Cohort 2.

^bWelch’s test for equality of variances was found to be violated. Due to this violated assumption, a *t* statistic not assuming homogeneity of variance was computed.

Similarly, at the teacher level, there were no statistically significant differences between

the treatment and control group teachers in the analytic sample at baseline on any of the demographic variables. See Table 2.

Table 2. Teacher Demographic Characteristics (2016–2018 School Years)

	Total Analytic Sample (N=123)			
	TSG (n = 59)		Control (n = 64)	
	%	%	χ^2 (df)	p
Gender			.85 (1)	.356
Female	64.41	56.25		
Education Level			.42 (2)	.812
Bachelors	64.41	60.94		
Masters	30.51	31.25		
Beyond MA	5.08	7.81		
Teaching Experience	Mean (SD)	Mean (SD)	t (df)	p
Total years of classroom teaching	10.37 (6.95)	9.94 (8.06)	.32 (121)	.750
Years teaching in Grade 8	5.90 (5.10)	5.61 (6.50)	.27 (121)	.786

In contrast, at the student level, treatment and control students differed significantly on one of the two pretest vocabulary measures at baseline; however, the difference is within the acceptable covariance adjustment range specified by the What Works Clearinghouse (What Works Clearinghouse [WWC], 2014). See Table 3.

Table 3. Baseline Characteristics of Student Analytic Sample (2016–2018 School Years)

Pretest	TSG n	Control n	Total Sample N	TSG Mean (SD)	Control Mean (SD)	t (df)	Hedges' g	p
Academic Vocabulary Test for 8 th Graders	2,061	2,018	4,079	21.68 (6.32)	22.11 (6.97)	-2.05 (4,078)	-.064	.041
Academic Word Mastery	2,004	1,980	3,984	21.31 (4.92)	21.37 (5.04)	-0.38 (3,983)	-.012	.707

Note: Sample size varies for each pretest because the analysis is based on all students with available data for that pretest measure.

At the student level, there were statistically significant differences between the treatment and control group students on ELL status and three race/ethnicity categories. See Table 4.

Table 4. Student Demographic Characteristics (2016–2018 School Years)

	Total Analytic Sample					
	TSG		Control		χ^2 (df)	p
	n	(%)	n	(%)		
Gender	1,977		1,966		1.50 (1)	.221
Female	959	48.51	992	50.46		
Male	1,018	51.49	974	49.54		
Race/Ethnicity ^a	1,964		2,004			
African American	281	14.31	268	13.37	0.73 (1)	.394

American Indian/Alaska Native ^b	181	9.22	133	6.64	9.08 (1)	.003
Asian	38	1.93	44	2.20	0.33 (1)	.564
Hispanic/Latino	1,358	69.14	1,513	75.50	20.02 (1)	< .001
Multiracial	748	37.89	740	36.82	0.49 (1)	.482
Native Hawaiian/Pacific Islander	3	0.15	6	0.30		.508 ^c
White	870	44.30	774	38.62	13.16 (1)	< .001
EL Status	1,502		1,785		27.77 (2)	< .001
Currently classified as EL	489	32.56	470	26.33		
Formerly classified as EL	315	20.97	502	28.12		
Not EL	698	46.47	813	45.55		

Note: Demographic data were not available for every student in the study. Pearson's chi-square test of independence (χ^2) was computed for students with demographic data, excluding honors students and students with IEPs/504 from both conditions.

^aPercentages for this category do not sum to 100 because multiple categories were indicated for many students.

^bThe sample size for the American Indian/Alaska Native category is 1,963 *TSG* and 2,004 *Control*.

^cSome of the expected value counts for the Native Hawaiian/Pacific Islander category were less than 5. Due to this violated assumption, Fisher's exact test was computed ($p = .508$).

Implementation of the TSG Professional Development Program

The *TSG* PD program was implemented in 25 Treatment schools. District personnel in conjunction with members of the research team selected school/district employees (district curriculum or PD specialist, classroom teachers) as the facilitators to implement the *TSG* PD sessions at the treatment schools. Coaches (IRG staff) provided on-going support to the facilitators based on audiotapes of their sessions.

The coaches listened to the audiotapes and provided detailed feedback to the facilitators in writing. They also had brief calls with the facilitators, highlighting key areas that went well and areas that could be improved for the next session. Coaches noted a wide range in the skill sets of the facilitators. Thus, while most implemented the program with fidelity, some facilitators required a good deal more support and guidance than others.

Procedural fidelity was assessed for Session 4 and Session 6 at each of the 25 implementation sites using fidelity checklists. The mean procedural fidelity for Session 4 was 84% (median = 86%) and for Session 6 was 76% (median = 80%). The quality of implementation at each site was assessed using a 5-point Likert scale for all 10 PD sessions. The mean overall quality rating was 4.0 (range = 3.8 to 4.7).

Impact of the TSG Professional Development Program on Teachers

We examined the impact of the *TSG* PD program on observed teaching practice in vocabulary, as measured by two *Observation Measure for Vocabulary Instruction (OMVI)* scales: *Teacher-Directed Vocabulary Instruction* and *Interactive Vocabulary Instruction*. We used a two-level random effects model, with teachers at level 1 and schools at level 2. The model tested the difference between conditions as a fixed effect.

Given various assumptions for covariates and multilevel models, we performed a sensitivity analysis that varied the covariates and random effects. First, as covariates, we included (a) no covariates; (b) years of teaching experience and teachers' education level (master's vs. bachelor's) at the teacher level; and (c) the teacher-level covariates plus the percent of students

classified as limited English proficient, percent with economic disadvantage, and percent White (nonminority) at the school level. We did not expect covariates to meaningfully affect the results due to the randomized design and their location on the causal pathway from condition to teaching behavior (Spector & Brannick, 2011). The results were highly consistent across models, varying by only a marginal degree.

Second, we considered three variations in random effects: (a) only observations and schools as random effects; (b) observations nested within teachers as subunits within schools; and (c) observations nested within teachers as subunits within schools and within random blocks. The blocks were created for randomization. Again, we expected minimal differences in the parameter estimates for condition among the three models. Murray, Hannan, and Baker (1996) and Zhu, Jacob, Bloom, and Xu (2012) have shown no change in Type I error rates when a model includes or excludes subunits. In addition, ignoring blocking, “though not faithful to the actual design used, does not appear to cause problems with either the Type I or Type II error rate” (Dong & Lipsey, 2010, slide 20). For example, Raudenbush and Sadoff (2008) demonstrated that blocks do not influence the variance of the treatment estimator without substantial treatment variability. The results were highly consistent across models, varying by only a marginal degree; in particular, neither intercepts nor treatment effects varied at the block level. See Table 5.

Table 5. Teacher-level Impacts Estimated with a Mixed-Model Analysis of Variance

		Teacher Explanations (54 df)	Student Practice (54 df)	OMVI Total (54 df)
Fixed Effects	Intercept	1.81**** (.41)	4.96**** (.81)	2.97**** (.24)
	Condition	2.72**** (.60)	3.82** (1.18)	1.44**** (.35)
Variances	School (Intercept)	3.65*** (.96)	13.60*** (3.72)	1.16*** (.35)
	Residual	4.90**** (.51)	21.00**** (2.16)	2.11**** (.22)
ICC		.423	.393	.354
Hedges' g	Condition	0.926	0.650	0.809
p-values	Condition	<.0001	.0021	.0001

****Significant at $p < .0001$; ***significant at $p < .001$; **significant at $p < .01$.

Impact of the TSG Professional Development Program on Students

Impact analyses for student level outcomes were based on a multilevel random effects model. Student level impacts were determined using a three-level model: students nested in teachers nested in schools. Two pretests Word Mastery and Academic Vocabulary Pretest were used as covariates. See Table 6 for impacts on the full sample and Table 7 for impacts on the ELL sub-sample.

Table 6. Student-level Impacts

	Academic Vocabulary Test for 8 th Graders (52 df)	Assessment of Social Studies Knowledge (ASK): Content Knowledge

		(52 df)	
Fixed Effects	Intercept	4.77**** (.38)	0.18 (.50)
	Condition	0.12 (.26)	0.18 (.46)
Variances	School (Intercept)	0.34 (.22)	1.71** (.63)
	Teacher (Intercept)	0.50* (.22)	1.38*** (.41)
	Residual	16.50**** (.41)	20.24**** (.50)
ICC		.020	.078
Hedges' g	Condition	0.018	0.028
p-values	Condition	.6519	.6983

****Significant at $p < .0001$; ***significant at $p < .001$; **significant at $p < .01$; *significant at $p < .05$.

Table 7. Impacts for the Sub-sample of ELL Students

		<i>Academic Vocabulary Test for 8th Graders (52 df)</i>	<i>Assessment of Social Studies Knowledge (ASK): Content Knowledge (52 df)</i>
Fixed Effects	Intercept	5.03**** (.48)	1.18 (.61)
	Condition	0.13 (.31)	0.44 (.54)
Variances	School (Intercept)	0.31 (.29)	1.96* (.89)
	Teacher (Intercept)	0.50 (.35)	1.97** (.72)
	Residual	16.72**** (.64)	18.79**** (.71)
ICC		.018	.094
Hedges' g	Condition	0.020	0.071
p-values	Condition	.6671	.4199

****Significant at $p < .0001$; ***significant at $p < .001$; **significant at $p < .01$.

Measure of Teacher Perceptions of Professional Culture

Two measures were used to gather data on teacher perceptions. The first, Nature of Professional Development, measures teachers' perceptions of how PD has influenced their teaching and the extent to which PD activities and topics are coordinated (coefficient alpha of .76). It was adapted from the University of Chicago Consortium on School Research (2007) survey measure. A researcher-developed measure, Professional Appraisal of TSG Survey, was used for purposes of formative evaluation. Thus, it was administered only to teachers in the experimental condition. See Tables 8 and 9.

Table 8. Teacher Satisfaction Survey – Part 1: Overall PD Experiences

Item	Percent of Teachers Responding				Percent of Teachers Responding			
	<i>TSG</i>				Control			
	Strongly disagree	Disagree	Agree	Strongly agree	Strongly disagree	Disagree	Agree	Strongly agree
1. Overall, my professional development experiences this year have:								
a. Been sustained and coherently focused, rather than short-term and unrelated.	2	15	64	20	0	14	66	20
b. Included enough time to think carefully about, try, and evaluate new ideas.	2	28	52	18	3	17	65	15
c. Been closely connected to my school's improvement plan.	3	13	62	20	0	8	60	32
d. Included opportunities to work productively with colleagues in my school.	2	11	54	31	2	14	52	32
e. Included opportunities to work productively with teachers from other schools.	34	31	21	13	12	31	43	12
2. How much do you agree with the following statements?								
a. Most of what I learn in professional development addresses the needs of the students in my classroom.	2	13	61	25	3	9	66	22
b. Most professional development topics are offered in the school once and not followed up.	3	41	46	10	9	54	32	5
c. Teachers are left completely on their own to seek out professional development.	25	59	11	3	45	38	6	11
d. Teachers in this school trust each other.	0	11	75	13	2	12	49	37
e. It's OK in this school to discuss feelings, worries, and frustrations with other teachers.	3	8	72	16	0	9	62	29
f. Teachers respect other teachers who take the lead in school improvement efforts.	0	5	69	26	0	8	52	38
g. Teachers in this school respect those colleagues who are expert at their craft.	2	7	59	33	2	5	46	48

Note. Some percentages do not sum to 100 because some teachers did not respond to all survey items

^aA small percentage of teachers answered with a range (i.e., disagree–agree) or halfway between two responses.

Table 9. Teacher Satisfaction Survey – Part 2: TSG Vocabulary PD Experiences

Item	Percent of Treatment Teachers Responding			
	Not at all helpful	Somewhat helpful	Helpful	Very helpful
1. Overall, in terms of assisting you to teach vocabulary to eighth graders, how helpful did you find the <i>TSG</i> Vocabulary PD program?	3	20	59	18
2. How beneficial is the <i>TSG</i> Vocabulary PD compared with other professional development activities you have attended?	Less beneficial	Somewhat less beneficial	Somewhat more beneficial	More beneficial
	3	11	59	26
3. How much do you agree with the following statements?	Strongly disagree	Disagree	Agree	Strongly agree
a. The information presented in the <i>TSG</i> Vocabulary PD was directly relevant to teaching and learning in my classroom.	2	7	70	21
b. The ideas presented in the <i>TSG</i> Vocabulary PD were easy to put into practice.	3	23	57	16
c. The <i>TSG</i> Vocabulary PD increased my knowledge of how I can teach vocabulary in my classroom.	0	7	54	39
d. I was provided with help during the <i>TSG</i> Vocabulary PD sessions if I was confused.	0	7	51	43
e. I learned different ideas in the <i>TSG</i> Vocabulary PD than I did in other professional developments I have attended in vocabulary.	5	8	56	31
f. My teaching skills in vocabulary have improved as a result of participating in the <i>TSG</i> Vocabulary PD.	2	10	59	30
g. The <i>TSG</i> Vocabulary PD material was presented clearly.	0	10	54	36
h. In the future, I plan to use the vocabulary strategies I learned in the <i>TSG</i> Vocabulary PD.	0	8	59	33
i. My students have responded positively to the vocabulary instruction.	0	10	70	18
j. I felt comfortable sharing my ideas and concerns during <i>TSG</i> Vocabulary PD sessions.	3	3	59	34
k. My students have benefitted from the vocabulary instruction.	0	8	70	20
l. Attending the <i>TSG</i> Vocabulary PD was a good use of my time.	7	15	61	18
4. How often did you implement the skills/ideas presented in the <i>TSG</i> Vocabulary PD?	Rarely	Sometimes	Most of the time	All of the time
	2	43	34	20
5. How useful were the following sessions of the <i>TSG</i> Vocabulary PD?	Not at all useful	Somewhat useful	Useful	Very useful

Item	Percent of Treatment Teachers Responding				
	Most helpful	2nd most helpful	3rd most helpful	4th most helpful	Least helpful
a. <i>Session 1: Words in Context</i>	3	20	44	33	
b. <i>Session 2: Selecting Words to Teach</i>	5	18	39	38	
c. <i>Session 3: Student Friendly Definitions</i>	0	18	34	48	
d. <i>Session 4: Examples, Non-examples, & Concrete Representations</i>	3	20	46	31	
e. <i>Session 5: Activities to Promote Word Learning</i>	0	16	49	33	
f. <i>Session 6: Cumulative Review</i>	2	25	48	23	
g. <i>Session 7: Using Context to Determine Word Meanings</i>	2	17	47	32	
h. <i>Session 8: Cumulative Review</i>	3	25	47	22	
i. <i>Session 9: Cumulative Review</i>	5	27	46	20	
6. Please rank the features of the <i>TSG</i> from Most Helpful to Least Helpful.					
a. <i>Debrief: Debriefed experiences in applying the research-based strategies to my teaching.</i>	16	8	23	16	36
b. <i>Discuss the Focus Research Concept: Discussed the research addressed in the readings.</i>	5	15	18	39	23
c. <i>Compare Research with Practice: Reviewed an upcoming lesson and discussed how it does or does not reflect the research principles discussed in the reading.</i>	8	11	38	21	21
d. <i>Plan Collaboratively: Designed lessons that incorporate research concepts.</i>	59	23	10	5	3
e. <i>Assignment: Teach lessons that you developed.</i>	25	48	11	10	7
7. Do you plan to use the vocabulary skills you learned in the <i>TSG</i> Vocabulary PD sessions next school year?	Yes	No	Undecided		
	95	3	2		
8. If a <i>TSG</i> were offered again at your school on another topic, would you volunteer to be part of it?	Definitely not volunteer	Might volunteer	Probably volunteer	Definitely volunteer	
	21	43	31	5	

Note. Some percentages do not sum to 100 because some teachers did not respond to all survey items.

Financial Cost of the *TSG* PD Program

The monetary cost of implementing the *TSG* PD program is reasonable and affordable. For school districts using literacy personnel already trained by our research team during our study, the cost would only be the *TSG* facilitator's guide currently available at \$35 per copy from IRG. A substitute teacher may also be required if the sessions are held during classroom hours, although we recommend conducting the sessions outside of classroom hours.

Schools or districts not previously in our study that would like to implement the intervention would need to consider the following costs:

Items	Cost
-------	------

Initial training of facilitators	This initial 2-day training will be provided by IRG trainers. The training can accommodate multiple facilitators.	\$3,500 per training
	Cost of substitute teachers.	Based on local costs
On-going support and coaching for facilitators over the duration of the PD (optional)	Coaching and support provided by IRG coaches. This cost includes labor and attendance of initial facilitator training.	\$4,000 per facilitator
	Coaching and support provided by local school district staff. Labor hours needed = 30 hours.	Based on local costs
TSG facilitator's guide	This is available from IRG.	\$35 per copy

All fees are based on reimbursement of our institute's current costs as of 2020 and are subject to change in accordance with rising costs. If the school district is outside of the Southern California area, travel cost and applicable overhead rates would also be charged.

C. What opportunities for training and professional development has the project provided?

Teachers in the treatment schools participated in the *TSG* PD program. The *TSG* sessions focused on providing in-depth vocabulary instruction for the words that appear in the American History texts and lessons. Teachers used their own curriculum and worked with grade level colleagues in developing lessons that include evidence-based vocabulary instruction.

Control schools were offered the option of our staff providing instruction in the *TSG* PD at the end of the study.

D. How were the results disseminated to communities of interest?

We presented at two national conferences on the *TSG* PD model (the 2018 and 2020 Council for Exceptional Children conferences). We plan to send an executive summary of the findings to each school that participated in the study. Given the fallout from COVID-19 and the resulting added burden on schools, we delayed informing the schools about the findings. We are also working on an article for publication in a peer-reviewed journal. We plan to disseminate the findings at conferences and inform the schools in the near future.

E. What do you plan to do during the next reporting period to accomplish the goals and objectives?

This is the final reporting period.

II. Products

A. Publications, conference papers, and presentations

We presented on the *TSG* PD model at two conferences. We presented at the Council for Exceptional Children conference in Tampa, Florida in February of 2018. The title of the presentation was: *The Vocabulary Conundrum: Which Words Do I Teach?* The abstract was: In this session participants will learn and practice a research-based procedure used in Teacher Study Groups for effectively and efficiently selecting vocabulary from content area textbooks.

We also presented at the Council for Exceptional Children conference in Portland, Oregon in February of 2020. The focus of the presentation was to describe the *TSG PD* program that was provided to the eighth-grade American History teachers in this study. The title of the presentation is: *The Vocabulary Conundrum: The Secret to Selecting Academic Vocabulary*.

B. Web site(s) or other Internet site(s)

Nothing to report; not applicable to the current study.

C. Technologies or techniques

IRG utilized several technologies in the performance of the study. A description and explanation of how they were used for research activities is provided below.

1. Student Vocabulary tests were scored in-house by IRG staff using a Scantron scanner and scoring software.
2. A digital audio recorder was used to record each of the *TSG* sessions. The audio recordings were used to assess implementation fidelity.
3. The *TSG* facilitator uploaded the audio recordings to Dropbox, where the coach was able to access them. Dropbox is an online backup facility which stores data on a secure server using Cloud Storage so that files can be shared with others using file synchronization.
4. SurveyMonkey, an online survey software and questionnaire tool, was used to administer professional development surveys to teachers. Participants completed their surveys online each month. The raw data were downloaded by IRG staff into excel files.
5. Data entry was completed using Microsoft Excel.
6. A randomization algorithm was developed using Microsoft Excel to randomly assign the two pilot schools to treatment and control conditions.

D. Inventions, patent applications, and/or licenses

Nothing to report; not applicable to the current study.

E. Other products

1. Updated *Observation Measure for Vocabulary Instruction (OMVI)*: The *OMVI* was fine-tuned to accommodate observations in classrooms with English learners.
2. Updated *OMVI* training for observers: IRG research team videotaped four middle school history teachers' vocabulary instruction. Selected clips from these videos were master coded and incorporated into the observation training.
3. Updated facilitator training: The research team modified the selections used in the facilitator training from first grade to selections from eighth-grade textbooks.
4. Updated *TSG PD*: The *TSG PD* program was modified for use with eighth-grade American History teachers. Specifically, Session 8 was deleted and substituted with a third cumulative review session.

- Academic Vocabulary Test for eighth graders: The research team developed an academic vocabulary measure for eighth graders. This measure will be used as an exploratory measure.

III. Participants and Other Collaborating Organizations

A. What individuals have worked on the project?

The project staff is led by the PD/PI, Russell Gersten, and two co-PIs, Madhavi Jayanthi and Joseph Dimino. The support staff include Sr. Research Associates, Research Associates, Research Assistants, and a team of field personnel including observers, testers, and coaches.

No personnel received additional funding support from other sources or conducted project activities in a foreign country. The individuals that worked on the project during the final reporting period and the nearest person months they had contributed are listed below.

Staff Name	Project Role	Nearest Person Month Worked*	Funding Support	Foreign Country
Russell Gersten	Director/PI	1	No additional funding support is provided	Did not collaborate with the individual in a foreign country
Madhavi Jayanthi	Co-PI	4		
Joseph Dimino	Co-PI	3		
Mary Jo Taylor	Sr. Research Associate	1		
Rebecca Newman-Gonchar	Sr. Research Associate	0		
Robin Schumacher	Research Associate	0		
Kelly Haymond	Research Associate	1		
Samantha Spallone	Research Associate	3		
Sarah Krowka	Research Associate	0		
Pamela Foremski	Research Assistant	1		
Christopher Tran	Research Assistant	0		
John Huynh	Business Manager	0		
Leslie Jensen	Observer/Tester/Coach	0		
Deborah Lewis	Observer/Tester/Coach	0		

*Individuals that worked more than 0 hours but less than half a month (85 hours) are listed with 0 nearest person month worked.

B. Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period?

There were no changes to the active support of the PD/PI or co-PIs since the last reporting period

C. What other organizations have been involved as partners?

There have been no partner organizations that were involved in the project.

D. Have other collaborators or contacts been involved?

None. Not applicable to the current study.

IV. Impact

A. What is the impact on the development of the principal discipline(s) of the project?

In the past decade, Gersten and colleagues at Instructional Research Group (IRG) examined the effectiveness of the *TSG* PD program in the areas of comprehension and vocabulary. The purpose of the current study was to extend the research by assessing the impact of the *TSG* PD program in academic vocabulary on (a) observed teaching practice, and (b) academic vocabulary outcomes of English learners when implemented with *eighth-grade teachers*, in the context of their social studies curriculum (i.e., American History).

Expected results from this study will provide evidence of the *TSG* PD program effectiveness in improving the academic vocabulary of English learners—a recommended area of instructional focus in the recently released IES Practice Guide on English learners (Baker et al., 2014). The study will also provide evidence regarding the feasibility of implementing the professional development program at the middle school level, where school culture and teacher dynamics tend to be markedly different from that of an elementary setting.

Focusing on improving academic vocabulary instruction is especially important for English learners, who by definition enter school with limited knowledge of the English language and are especially weak in fluency with the formal language of academic disciplines (Lesaux et al., 2010; Scarcella, 2003; Snow et al., 2009). However, this approach is likely to be equally beneficial to students who are English proficient, but who come from low-income families and communities and have limited experiences in using formal Standard English (e.g., Hart & Risley, 2003). By the time they reach eighth grade, students who started off behind often demonstrate even larger gaps in academic vocabulary knowledge than they did in earlier grades (Biemiller & Slonim, 2001; Lesaux et al., 2010). This is precisely the reason eighth grade teachers need to provide particularly robust vocabulary instruction to these students.

The studies on *TSGs* as a means of PD are cited in several recent meta-analyses on professional development and seem to fit the general pattern of leading to statistically significant impacts in aspects of teacher classroom performance, but not on measures of student learning. These findings also are being used as IES develops its practice guide on reading interventions for students in grades 4–9 in the area of vocabulary. Future dissemination work will highlight the need for teachers to also receive specific curriculum material or supplements that they can use in their classroom to possibly lead to positive outcomes. Future dissemination work will also focus on the difficulty in measuring the impact of vocabulary instruction on vocabulary learning in large-scale studies, where different teachers teach different vocabulary words, and standardized measures seem invariably insensitive to these measures.

B. What was the impact on other disciplines?

This project did not make any impact on other disciplines.

C. What was the impact on the development of human resources?

The purpose of this study is to provide professional development in academic vocabulary for eighth-grade American History teachers. As a result of this year long, job-embedded professional development program, we anticipate teachers will learn effective ways to teach academic vocabulary to their students. The professional development program will help in building teacher capacity to effectively teach academic vocabulary words in an engaging, evidence-based fashion to their students. Teachers participating in the professional development program will learn two basic skills: (a) how to decide which words to teach to their students based on certain criteria (e.g., conceptually central to understanding the text, student background knowledge) and (b) how to teach these words in an engaging manner (e.g., student friendly definitions, examples) such that these words become part of students' listening, speaking, reading and writing vocabulary.

In addition to the benefits for teachers, we anticipate a positive impact for the facilitators. The facilitator training will help build capacity of professional development staff to provide trainings in teaching academic vocabulary words effectively at the middle school level.

Finally, this research might hopefully lead to development of curriculum supplements so that teachers can use them in eighth-grade American History. The principles behind the *TSG* model made sense to most teachers, and many indicated that if they possessed such materials, they would be able to effectively teach more vocabulary words and important concepts involved in truly understanding the meaning of words such as "compromise", "abolitionist", "divisive."

D. What was the impact on teaching and educational experiences?

The observed teaching practices of teachers in the *TSG* condition were significantly better than those in the control condition, indicating that the PD has the potential to impact the vocabulary instructional practices in social studies classrooms. Survey data indicated that a majority of the teachers found the PD to be valuable, and teachers at some sites noted in informal communications that they saw improved student understanding of vocabulary concepts.

E. What was the impact on physical, institutional, and information resources that form infrastructure?

This project did not make any impact on physical, institutional, or information resources from infrastructure.

F. What is the impact on technology transfer?

This project did not make any impact on technology transfer.

G. What is the impact on society beyond science and technology?

None. Not applicable to the current study.

H. What percentage of the award's budget was spent in foreign country(ies)?

None. Not applicable to the current study.

V. Changes/Problems

A. Changes in approach and reasons for change

Nothing to report.

B. Actual or anticipated problems or delays and actions or plans to resolve them

Nothing to report.

C. Changes that have a significant impact on expenditures

Nothing to report.

D. Significant changes in use or care of human subjects, vertebrate animals, and/or biohazards

Nothing to report.

E. Change of primary performance site location from that originally proposed

Nothing to report.

VI. Special Reporting Requirements

Nothing to report.

VII. Budgetary Information

During the reporting period of March 1, 2019 – June 30, 2020, the remaining authorized funds were fully expended. The funds supported the remaining activities: (a) data collection and analyses; (b) draft of Final Report; and (c) attend annual PI meeting and dissemination events. Labor cost and the associated fringe cost constitute most of the funds that were spent. These dollars paid for the wages and benefits of key personnel and other personnel to conduct the remaining tasks. Travel and other direct cost represent the other expenditures. Before COVID-19 ended all unnecessary travel, those dollars were spent for staff to attend dissemination events.

The table below breaks down the spending during the final reporting period and over the course of the entire project. During the final budget period, we expended \$376,882. Over the course of the project, we spent \$3,475,094, resulting in no remaining balance.

	Authorized Funding for the Entire Project Period (Total Award)	Draw Downs during Current Budget Period (Since 2/28/2019)	Spending for the Entire Project Period (Total Drawdowns)	Unexpended Funds (Available Balance)
Labor Cost				
Key Personnel	\$ 764,700	\$ 140,126	\$ 703,317	\$ 61,382
Other Personnel	\$ 925,899	\$ 49,109	\$ 914,990	\$ 10,909
Fringe Cost	\$ 638,775	\$ 91,131	\$ 599,375	\$ 39,400
Consultants	\$ 28,500	\$ -	\$ 15,759	\$ 12,741
Travel	\$ 75,097	\$ 3,883	\$ 293,619	\$ (218,521)
Materials	\$ 30,505	\$ 43	\$ 25,237	\$ 5,268
Stipend	\$ 79,900	\$ -	\$ 62,850	\$ 17,050
Study Compensation	\$ 38,775	\$ -	\$ 18,700	\$ 20,075
ODC	\$ 40,563	\$ 1,195	\$ 44,130	\$ (3,568)
Total Direct Cost	\$ 2,622,713	\$ 285,487	\$ 2,677,977	\$ (55,264)
Indirect Cost	\$ 852,382	\$ 91,394	\$ 797,117	\$ 55,265
Total Direct and Indirect	\$ 3,475,095	\$ 376,882	\$ 3,475,094	\$ 1

The breakdown of funds expended in the table above shows over-spending on travel over the course of the project. As previously reported, this was due to more participating schools being located outside IRG’s local area than had been originally budgeted for, requiring field staff to travel out-of-town to implement the study. To offset the over-spending on field staff and travel, office staff were used more efficiently to reduce labor and fringe costs. Overall, there is no impact on the total budget or our ability to achieve project goals.

VIII. Project Outcomes

Structured Abstract

Setting

This project took place in eighth-grade classrooms in middle schools in urban school districts in Arizona, California, Colorado, Florida, Georgia, and Texas.

Sample

The sample for the study consisted of 123 eighth-grade teachers of American History and their 4,227 students in 54 middle schools, with the majority of schools having at least 10% of the student body being English language learners. Analysis was also conducted using a subsample of 1,776 ELL students.

Program

The fully developed *Teacher Study Group (TSG)* professional development (PD) program is a concentrated PD effort designed to improve teaching practice in order to produce increases in

student academic vocabulary outcomes. Providing rich, in-depth vocabulary instruction, the *TSG PD* has been designed for teachers in high poverty, Title I schools, with a wide range of learners, including linguistically diverse learners from various ethnic backgrounds. The *TSG PD* consists of 10 interactive sessions, which take place twice each month from October to March. Each session lasts 75 minutes. During the sessions, a five-phase recursive process is instituted: (1) Debrief Previous Application of the Research, (2) Discuss the Focus Research Concept, (3) Compare the Focus Research Concept with Practice, (4) Plan Collaboratively, and (5) Assignment. The goal of the *TSG PD* is to enhance instruction by helping teachers integrate research-based instructional strategies into their existing curriculum.

Research Design and Methods

The study used a multi-site cluster randomized trial, with schools randomly assigned to conditions within districts. There were two conditions: receiving the *TSG PD* or not receiving the *TSG PD*. The research team administered pretest measures 3–4 weeks prior to the beginning of the *TSG PD*. Researchers collected various measures during the program period and collected posttest measures 3–4 weeks after the program finished.

Control Condition

Teachers in the control condition did not receive the *TSG PD*, but they may have, per business-as-usual practices, received professional development from some other source during the study period.

Key Measures

Among the many instruments used in the study, researchers used the *Observation Measure for Vocabulary Instruction* to measure observed teaching practice, and the *Academic Vocabulary Test for 8th Graders* and the *Assessment of Social Studies Knowledge (ASK): Content Knowledge* to measure student academic vocabulary and content knowledge, respectively.

Data Analytic Strategy

Researchers used multilevel modeling to address the primary research questions regarding teacher outcomes and student outcomes.

Cumulative Outcomes

Teachers who participated in the *TSG PD* demonstrated greater use of effective vocabulary instructional practices than teachers who did not participate, and the difference was statistically significant. The *TSG PD* program did not demonstrate any impact on student academic vocabulary or content knowledge. Students of teachers who participated in the *TSG PD* performed at almost the same level as students of teachers who did not participate; the difference was not statistically significant.