Evaluation of a Project-based Science Intervention for High School Chemistry and Physics

Quinton Baker, Christopher Klager, Barbara Schneider, & Joseph Krajcik
Michigan State University

Background/Context

The Next Generation Science Standards (NGSS; NGSS Lead States, 2013) have been recently adopted in many states, and mark a change from content-focused instruction to three-dimensional learning that incorporates disciplinary core ideas with scientific and engineering practices and crosscutting science concepts (NRC, 2012). Project-based learning (PBL) is an instructional strategy that aligns well with the intentions of the NGSS through its focus on explaining phenomena using scientific practices. Previous research on PBL interventions in science classrooms have shown that it can increase student science learning (e.g. Harris et al., 2015), although it has not been rigorously tested in high school chemistry or physics.

Purpose/Objective/Research Question

This study is a cluster-randomized trial that tests PBL units designed for high school chemistry and physics classes along with an associated professional development system. Our research questions are: 1) Does the program improve student science learning on a three-dimensional science assessment compared to the control group? and 2) Is the effect of the intervention moderated by student-level characteristics such as gender and race/ethnicity? We hypothesize that this PBL curriculum will improve science learning as measured on a three-dimensional science post-test compared to a "business-as-usual" control group.

Setting

This cluster-randomized trial took place in high school chemistry and physics classrooms in one Midwestern and one Western state.

Population/Participants/Subjects

High school chemistry and physics students from 60 schools participated in the study. The analytic sample consists of 3320 students.

Intervention/Program/Practice

Treatment teachers attended a three-day professional development workshop, led by trained project staff, which focused on the NGSS, PBL, and the curriculum materials. The curriculum materials included lesson plans for three, three-week units in either chemistry or physics. These units were designed using the principals of project-based science (e.g. Krajcik & Shin, 2014) and emphasized students' use of scientific practices to explain natural phenomena. Teachers were also provided with materials for completing lab activities and unit assessments aligned with unit learning goals. As teachers implemented the units, group teleconference meetings were held where teachers could share ideas and get assistance from trained project staff.

Research Design

This study was a cluster-randomized trial. Schools were assigned to treatment and "business-as-usual" control conditions (all schools had a .5 probability of being assigned to treatment or control), resulting in all chemistry and physics teachers in a single school being assigned to the same condition to prevent spillover due to sharing materials and knowledge. Treatment teachers received all elements of the treatment as described earlier, while control teachers received two-days of professional development on the NGSS. Baseline characteristics of the treatment and control conditions are presented in Table 1. We find small differences between the treatment and control analytic samples in standardized pretest score (-.08 standard deviations) and by proportion of the sample in each race/ethnicity category. We control for these differences in the analytic model.

Data Collection and Analysis

At the beginning of the school year, students were given a science content pretest, composed of publicly available NAEP items, in order to establish baseline equivalence (U.S. Department of Education, 2017). Students and teachers also completed background surveys that collected information about demographic characteristics and orientation toward school and science. At the end of the school year, all students were administered a post-test which was designed by a state department of education for use as their state's high school physical science assessment. This outcome measure was designed to cover the three-dimensions of science learning (NRC, 2012). Students and teachers were also given items from the background survey, to measure post-intervention levels of key constructs. Teachers also reported their use of features of project-based learning and scientific practices. Throughout the implementation of the intervention, teachers were observed during each instructional unit and rated by the observer on the fidelity of implementation of the curriculum.

A hierarchical linear model (HLM; Raudenbush & Bryk, 2002) was used to assess the difference between the treatment and control condition on science achievement, accounting for clustering that occurs as a result of assignment of schools to treatment (Bloom, 2005; Raudenbush, 1997). We estimate the following model:

$$PostTestScore_{ij} = \beta_0 + \beta_1 Treatment_j + \beta_2 PreTest_{ij} + \beta_3 SchoolMeanPretest_j + \beta_4 Chemistry_{ij} + \beta_5 Region_j + \beta_6 \boldsymbol{X_i} + e_{ij} + u_j$$

where $PostTestScore_{ij}$ is the post-test score for student i in school j. β_1 is the coefficient of interest, indicating the mean difference between the treatment and control groups on the science post-test. We control for individual and school-mean pre-test scores ($PreTest_{ij}$ and $SchoolMeanPreTest_j$), an indicator for chemistry versus physics ($Chemistry_{ij}$), an indicator for region where school $_j$ was located ($Region_j$), and include error terms at the student (e_{ij}) and school levels (u_j).

Variations of this model include interactions between the treatment indicator and the chemistry indicator to look for differential effects by subject area, and interactions between the treatment indicator and demographic characteristics such as race and gender (in $\boldsymbol{X_i}$) to look for heterogeneous treatment effects. Measures of fidelity of implementation based on the observer ratings and teacher-reported use of the curriculum were used as mediators of the treatment effect.

Findings/Results

Table 2 contains the results of the multi-level regressions. Overall, we find that the treatment condition outscored the control condition by .30 standard deviations on the science post-test. We find no differential impact when interacting the treatment indicator by science discipline, gender, or race except that it appears Asian students saw no gains from the treatment.

Conclusions

Our results indicate that the CESE system for science learning provided students in the treatment group with a substantively important increase in their performance on the three-dimensional science learning outcome compared to control students. This result shows the promise of PBL approaches for teaching high school chemistry and physics as teachers work to modify their instruction to align with the NGSS. Although this finding cannot be generalized outside of the population represented in this study, the study sample was diverse both socio-economically and racially, indicating that this intervention may hold promise for a wide variety of students. The professional development and lab materials for this study were costly, and it is yet to be seen how well such an intensive intervention can scale.

Table 1. Baseline characteristics of treatment and control groups

	Control (N = 1633)	Treatment (N = 1687)	Difference
Standardized Pretest	0.08	0.01	0.08*
Percentage Female	0.53	0.53	0.00
Percentage Asian	0.06	0.04	0.01
Percentage Black	0.08	0.12	0.04***
Percentage Hispanic	0.48	0.43	0.05*
Percentage White	0.33	0.31	0.02
Percentage Other Race	0.06	0.05	0.01

Note. Differences in baseline characteristics are calculated for the analytic sample.
* p < .05; ** p < .01; *** p < .001

Table 2. Multi-level regression of standardized science post-test scores

	Main Model	Female Interaction	Race Interactions
Treatment	0.30 (0.07)***	0.28 (0.06)***	0.28 (0.12)*
Pre-test	0.29 (0.03)***	0.28 (0.03)***	0.29 (0.03)***
School-level Pre-test	0.32 (0.09)***	0.32 (0.09)***	0.31 (0.09)***
Chemistry	0.28 (0.10)**	0.28 (0.10)**	0.29 (0.10)**
Female	-0.02 (0.03)	-0.05 (0.04)	-0.02 (0.03)
Race/Ethnicity			
Asian	-0.02 (0.08)	-0.02 (0.08)	0.10 (0.06)
Black	-0.26 (0.10)*	-0.26 (0.10)**	-0.37 (0.16)*
Hispanic	-0.20 (0.09)*	-0.20 (0.09)*	-0.21 (0.12)
Other Race	-0.20 (0.11)	-0.20 (0.11)	-0.30 (0.15)
Region Dummies	O	O	O
Interactions			
Treatment x Female		0.05 (0.07)	
Treatment x Asian			-0.28 (0.11)*
Treatment x Black			0.19 (0.17)
Treatment x Hispanic			0.02 (0.13)
Treatment x Other Race			0.23 (0.20)

Note. Robust standard errors are in parentheses. Pre-test and the outcomes (post-test) are standardized with mean equal to zero and standard deviation equal to one.
* p < .05; ** p < .01; *** p < .001

References

Bloom, H. (2005). *Randomizing groups to evaluate place-based programs*. New York, NY: MDRC.

Harris, C. J., Penuel, W. R., D'Angelo C. M., DeBarger, A. H., Gallagher, L. P., Kennedy, C. A., Cheng, B. H., & Krajcik, J. S. (2015). Impact of project-based curriculum materials on student learning in science: Results of a randomized controlled trial. *Journal of Research in Science Teaching, 52*(10), 1362-1385.

Krajcik, J. S., & Shin, N. (2014). Project-based learning. In R. K. Sawyer (Ed.), *Cambridge Handbook of the Learning Sciences* (2nd ed., pp. 275-297). New York, NY: Cambridge University Press.

National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: The National Academies Press. doi:10.17226/13165

NGSS Lead States. (2013). *Next generation science standards: For states, by states*. Washington, DC: The National Academies Press. https://doi.org/10.17226/18290

Raudenbush, S. W. (1997). Statistical analysis and optimal design for cluster randomized trials. *Psychological Methods, 2*(2), 173-185.

Raudenbush, S. W., & Bryk, A. (2002). *Hierarchical linear models* (2nd ed.). Thousand Oaks, CA: Sage Publications.

U.S. Department of Education (2017). *What Works Clearinghouse: Standards handbook version 4.0*. Washington, DC: IES.