Implementing a Problem-Based Learning Curriculum in a University-School Collaborative Project for Improving Middle School Science Education: Lessons from Year One on Fidelity of Implementation

Brian D. Gane, Ph.D.
bgane@gatech.edu
Georgia Institute of Technology

Cher C. Hendricks, Ph.D.
cher.hendricks@ceismc.gatech.edu
Georgia Institute of Technology

Purpose

Science Learning Integrating Design, Engineering, and Robotics (SLIDER) is a 5-year project funded by the National Science Foundation to design and implement a problem-based learning (PBL) curriculum that utilizes LEGO robotics to teach 8th grade physical science. The project involves collaboration among K-12 educators, university faculty, and educational specialists located in our university's center for K-12 STEM outreach. Our study takes place in three public schools across the state with diverse student populations. The project also utilizes graduate students from STEM fields who serve as teaching Fellows in the schools to support curriculum implementation. Our research questions can be categorized as either formative (e.g., What type of support is necessary to adequately prepare teachers to deliver the curriculum? How should learning be assessed?) or summative (e.g., In what ways does the curriculum impact student content knowledge and reasoning skills? In what ways does it impact engagement, motivation, attitude, and interest in STEM fields?).

In our first year, we worked with seven participating teachers to facilitate their use of PBL instruction, designed and pilot tested the first units of the PBL LEGO curriculum, and collected pilot data. In the coming year (Year 2), teachers will continue to implement PBL curriculum, including three units of the LEGO robotics curriculum. In Year 3 teachers will implement the full LEGO robotics curriculum.

At the end of Year 1, we identified a number of concerns related to curriculum development, professional development needs, teacher and administrator buy-in, and differences across school sites. These concerns can be framed within a larger issue: will the curriculum be implemented with fidelity across our school sites and within each classroom?

The purpose of this paper is to describe (1) Year 1 findings regarding the challenge of ensuring teachers faithfully implement our curriculum and (2) our plan to simultaneously measure and improve fidelity of implementation. Our long-term research goal is to assess the efficacy of our curriculum while using the interim fidelity data to inform curriculum development and professional development.

Theoretical Framework

Problem-based learning is a broad approach to instruction with the central tenet that students can learn by attempting to solve authentic, complex problems in their domain of study. In the process of solving problems, students learn relevant declarative knowledge, improve their skills using domain-specific problem solving methods, and develop metacognitive ability to direct and assess their own learning. PBL can be as effective in teaching declarative knowledge as "traditional" instruction, while improving students' engagement and motivation (Hmelo-Silver, Duncan, & Chinn, 2007; Kolodner et al., 2003).

Fidelity of Implementation. Although PBL methods have shown positive results in classroom studies, they are challenging for teachers to enact. Based on our colleagues' earlier research, we know implementing a PBL curriculum requires significant training and multiple years of practice. Moreover, teachers frequently adapt PBL curricula to meet classroom and administration demands. For instance, teachers might change (or create) assessments to encourage on-task behavior, create accountability, or to prepare for standardized testing (Pedersen, Arslanyilmaz, & Williams, 2009). Given that teachers, not the curriculum designers, are driving these adaptations (often without any input to/from the curriculum designers) these changes run the danger of becoming "lethal mutations" (Brown & Campione, 1996), undermining curriculum efficacy.

Adaptations are common when moving from design to implementation in a classroom. *Fidelity of implementation* (FOI) refers to the extent to which the implemented intervention matches the intent of the designers. In other words, when teachers are using the curriculum, how closely do their classroom practices follow the intent of the curriculum developers?

We adopt Century, Rudnick, and Reeman's (2010) theoretical framework for FOI, which is based on identifying *critical components* that are unique and central to the intent of the intervention. The critical components involve both *structural* (e.g., whether the curriculum sequence is delivered in the designed order) and *instructional* components (e.g., whether the student is engaged in critical thinking regarding the problem). Reform-based instruction, such as PBL, is sometimes implemented with structural components but without the associated instructional components (e.g., Borko, Stecher, Alonso, Moncure, & McClam, 2005). For example, a teacher might use a critical thinking activity but deliver it to students with a didactic approach, eliminating the cognitive engagement it was designed to foster.

During Year 1 we realized we would need to systematically measure FOI. Additionally, we determined the curriculum designers would need to work closely with teachers to increase the likelihood that critical components were maintained even after classroom adaptation occurs. In the following sections, we describe our data collection method, results, and major findings that lead us to these conclusions. We conclude by describing our plan for adapting Century and colleagues' (2010) FOI framework to our intervention.

Methods

Our project team includes learning theory and cognitive science specialists, a LEGO robotics educator, and an educational researcher. The project team includes a curriculum development team that is creating the PBL robotics curriculum that aligns with state performance standards for 8th grade physical science. During Year 1 the curriculum team developed the first unit; portions of it were pilot tested in two schools with five teachers.

Our project team also includes a research team that focuses on operationalizing variables, locating and creating measurement instruments, conducting field observations, and collecting data to inform curriculum development and measure student outcomes.

As the curriculum and research teams worked together during Year 1, they realized one major challenge would be capturing how teachers differentially implemented the curriculum in their classrooms. Because we had not previously considered how to explicitly measure FOI, we started by examining the existing data sources—many of which were informal—to begin to assess FOI and consider how to design the professional development necessary for teachers to implement the curriculum with fidelity.

For this paper, our method involved analyzing a subset of existing indicators (at the school, teacher, and student levels) that could better illustrate the issues we informally observed. In the following sections we describe data sources and report a subset of indicator data at each level, and we conclude by integrating these into an emerging picture of how our curriculum is being enacted in the classroom.

Data Sources

We relied on several sources of data during Year 1, including data on schools, teachers, and students (see Table 1). Data sources were (1) school and Department of Education (DOE) records (student and teacher demographics, student test scores), (2) teacher report (provided in informal conversations or email exchanges with teachers), (3) field observations of teachers implementing PBL curriculum, (4) Fellows' journal entries and interview responses, (5) teacher questionnaires about (a) their training, years experience, education level, types of students served, and whether they worked with co-teachers and (b) their experiences with their Fellows, (6) teachers' responses to the *Patterns of Adaptive Learning Scale* (PALS), which is used to measure teacher perceptions of school goal structures and teachers' goal-related approaches to instruction as well as personal teaching efficacy (Midgley et al., 2000), (7) teachers' responses on a survey that measures how much they value various PBL strategies and how much they actually engage in those strategies (based on Llewellyn's (2007) *Rubric for Becoming and Inquiry-Based Teacher*), and (8) concerns sheets completed by teachers during professional development.

Table 1. Sources of School, Teacher, and Student Data

Data about Schools	Data about Teachers	Data about Students
School records	Field observations	School/DOE records
Teacher report	Fellow journals/interviews	Teacher report
Field observations	Teacher questionnaires	Field observations
Fellow journals/interviews	PALS survey	Fellow journals/interviews
	PBL values/practices survey	
	Teacher concerns sheets	

Results and Conclusions

Schools. Our three participating schools differed in teacher workload and time for implementing the curriculum. Measures include contact hours, number of science courses taught, number of non-intervention courses taught, and total number of students. First, the schools differed in contact hours (see Table 2). School 3 had almost 40 *fewer* contact hours than Schools 1 and 2; in Year 2 it has 60 *more* contact hours than Schools 1 and 2. Second, two of our seven teachers taught both physical science and social studies. Both expressed difficulty in having sufficient planning time to teach two subjects. Third, teachers taught a different number of physical science classes per day (ranging from 3 to 6) and class size varied substantially. For example, one teacher taught 4 periods of physical science with a total of 58 students. Another taught 6 periods of science with a total of 145 students. Class size ranged from 14.5 students per period for one teacher to 26.4 for another teacher. These three factors affect implementation, for example, by constraining time to plan and time to deliver the curriculum.

Table 2. School Differences in Time for Implementing Curriculum

School	Instructional Days	Period Length (minutes)	Contact Hours
	Year 1		
1	180	55	165.0
2	180	55	165.0
3	170	45	127.5
	Year 2		
1	180	55	165.0
2	180	55	165.0
3	180	75	225.0

Teachers. Unsurprisingly, teachers differed along a variety of dimensions, but we identified several that would likely affect FOI. Of these, teacher efficacy and teachers' values and practices related to PBL are reported here. The PBL areas assessed were curriculum (e.g., *planning and designing student-centered curriculum*), lesson presentation (e.g., *creating units that begin with a highly motivating problem*), engagement (e.g., *engaging students in investigations, discourse, and reflection*), questioning (e.g., *using probing statements,*

prompts, and redirecting questions to solicit students' understanding), and assessment (e.g., *observing and assessing critical thinking skills*). Teachers responded on a 4-point scale for PBL questions and a 5-point scale for efficacy. As indicated in Table 3, there were consistent differences across teachers in terms of reporting more value for PBL strategies than in engaging in them, but differences between teachers were large in a number of areas. Teachers' belief in the value of PBL strategies and self-efficacy likely affects which PBL strategies they will engage in and the extent to which they will persevere when encountering difficulties.

Table 3. Teacher PBL and Self-Efficacy Responses

	PBL Values/Practices Survey										**PALS Survey**
	Curriculum		**Lesson Presentation**		**Engagement**		**Questioning**		**Assessment**		**Teaching Efficacy**
Teacher	Value	Practice	Value	Practice	Value	Practice	Value	Practice	Value	Practice	
1	3.7	2.7	3.7	2.7	3.3	4.0	2.8	3.8	3.8	3.0	3.3
2	3.8	2.5	3.7	2.5	3.5	3.0	3.8	3.2	3.6	3.2	3.7
3	4.0	2.3	4.0	2.2	4.0	2.4	4.0	2.8	4.0	2.8	3.6
4	4.0	2.5	3.2	2.2	3.5	2.5	3.3	2.0	2.4	1.6	3.0
5	4.0	3.2	3.7	2.8	4.0	3.0	3.8	2.8	3.6	3.0	3.9
6	3.6	2.0	3.3	1.3	4.0	2.0	4.0	2.0	3.6	2.0	2.4
7	4.0	2.3	3.8	1.8	3.8	2.7	4.0	3.2	4.0	2.6	4.0
Mean	3.9	2.5	3.6	2.2	3.7	2.8	3.7	2.8	3.6	2.6	3.4

Students. Students differed in content knowledge, academic preparation, and skill at taking standardized tests. Scores from our state's standardized test (CRCT) differed across schools in percentage of students "not meeting" and "exceeding" state standards in science (School 1: not met = 44%, exceeded = 21%; School 2: not met = 36%, exceeded = 9%; School 3: not met = 30%, exceeded = 10%). These data suggest substantial variability in science knowledge exists *within* classrooms. Some teachers were reluctant to stray from traditional curriculum and instructional strategies, fearing our curriculum would not adequately prepare students for the CRCT.

Additional FOI factors are outlined in Table 4.

Table 4. School, Teacher, and Student FOI Factors, Descriptions, and Indicators

Factors	Description	Current Indicator
School		
Benchmark Testing	Schools use different benchmark tests and are on a different scope and sequence (material is covered and tested at different times across sites).	Teacher report; Fellow journals; teacher concerns sheet
Planning time/use requirements	Schools require different activities during planning time across sites. At some sites, there has been very little planning (either individual or team planning) for curriculum delivery.	Teacher report; observations; Fellow journals
Administrator support	Teachers report administrators say they support the new curriculum but hold them to standards of the old curriculum. They also report that some administrator decisions are counter to what the program requires.	Teacher report; teacher concerns sheet
Grading requirements	Different schools and districts require teachers to use different grading policies (e.g., no grade below 50 or no 0 grades) and weightings for various assignments.	Teacher report; teacher concerns sheet
Contact Hours	Two schools were on a 180-day schedule with 55-minute science periods; the third was on a 170-day schedule with 45-minute science periods.	Teacher report; school records
Number of classes or preps	Some teachers had one class to prepare/ teach (physical science); others taught 2 different subjects. There was a wide range of number of periods taught and number of students.	Teacher report; observations; school records
Class makeup (SWD/co-teacher)	Some teachers were assigned more students with disabilities; some teachers also worked with co-teachers for all or part of the day.	Teacher report; observations; school records; teacher concerns sheet
Teacher		
Science content knowledge	Teachers varied in their physical science content knowledge.	Informal content assessment; teacher report; observations; Fellow journals and interviews
Value/use of PBL strategies	Teachers varied in how much they valued various PBL strategies and how much they engaged in these strategies.	PBL Values/Actions Survey; teacher report; observations; Fellow journals
Teaching efficacy	Teachers varied in their beliefs that they significantly contribute to students' academic success and are able to effectively teach all students.	*Patterns of Adaptive Learning Scale*

Classroom control	Teachers varied in their orientation to classroom control. Some were more willing to try new strategies as well as relinquish control (e.g., co-teach) with Fellows.	Fellow journals and interviews; Teacher surveys of work with Fellows.
Student		
Students with disabilities (SWD)	There were differences across schools and between teachers in numbers of SWD students served.	Teacher report; observations; school records
English language learners (ELL)	There were differences across schools in numbers of ELLs served.	Teacher report; observations; school records; Fellow journals and interviews
Socioeconomic status (SES)	There were differences across schools in student SES.	Teacher report; observations; school records
Achievement Scores (CRCT)	There were differences across schools in students' Criterion-Referenced Competency Test (CRCT) scores, the state's measure of Annual Yearly Progress.	Department of Education reports
Prior LEGO experience	Students at two schools had access to additional courses and after-school programs that used LEGO robotics.	Teacher report

These data capture an emerging picture of classroom differences motivating our need to assess FOI. As indicated in Table 4, we have identified multiple informal indicators, but they are not integrated within an overarching framework. This absence of a theoretical framework makes it difficult to aggregate the indicators into an *overall measure* of FOI. Moreover, it forestalls grouping individual indicators into conceptual classes that capture the type and source of fidelity.

Due to limitations of our existing indicators, we decided on a more principled approach. We propose adopting Century and colleagues' (2010) FOI framework, which provides several advantages: providing an underlying theoretical model on which to develop a formalized plan for measuring FOI, specifying high-level categories of implementation fidelity (i.e., critical components), and providing a multi-method suite of instruments that focus on different levels of observation.

Significance

In this paper we documented the kinds of challenges that occur when university-school collaborative programs are implemented that focus on curriculum studies. We focused on one challenge: fidelity of implementation. Multiple informal indicators provided preliminary data suggesting the need to measure FOI. Having identified this need, we propose a solution to adapt an existing FOI framework to our project. This involves identifying the critical components of our designed curriculum and revising various

instruments to assess FOI. While assessing FOI, we will work with teachers to ensure that the enacted curriculum retains the critical components of the intended curriculum.

Our experience also provides some "lessons learned" for other university-school collaborative researchers:

1. When working with teachers who are not also the curriculum developers, FOI should be considered from the beginning of the project.

2. Irrespective of teachers' stated willingness to adopt reform-based curriculum, a variety of issues will affect implementation (e.g., amount of class time, teachers' self-efficacy, students' content knowledge).

3. FOI issues might be exacerbated by using multiple schools, but they are in no way avoided by using a single school or a single teacher.

4. Although intervention researchers sometimes measure FOI (see O'Donell for a review), they often do not ground it within a theoretical framework. Adopting a framework allows one to characterize ways in which implementation changes and allows researchers finer-grained data collection.

References

Borko, H., Stecher, B.M., Alonzo, A.C., Moncure, S., & McClam, S. (2005) Artifact packages for characterizing classroom practice: A pilot study. *Educational Assessment, 10*, 73-104. DOI: 10.1207/s15326977ea1002_1

Brown, A.L., & Campione, J.C. (1996). Psychological theory and the design of innovative learning environments: On procedures, principles, and systems. In L. Schauble & R. Glaser (Eds.), *Innovations in learning: New environments for education* (pp. 289-325). Mahwah, NJ: Erlbaum.

Century, J., Rudnick, M., & Freeman, C. (2010). A framework for measuring fidelity of implementation: A foundation for shared language and accumulation of knowledge. *American Journal of Evaluation, 31*(2), 199-218. DOI: 10.1177/1098214010366173

Hmelo-Silver, C.E., Duncan, R.G., & Chinn, C.A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist, 42*, 99-107.

Kolodner J.L., Camp, P.J., Crismond, D., Fasse, B., Gray, J., Holbrook, J., Puntambekar, S., & Ryan, M. (2003). Problem-based learning meets case-based reasoning in the middle-school science classroom: Putting Learning by Design™ into practice. *Journal of the Learning Sciences, 12*, 495-547.

Llewellyn, D. (2007). *Inquire within: Implementing inquiry-based science standards in grades 3-8* (2nd ed.). Thousand Oaks, CA: Corwin Press.

Midgley, C., Maehr, M.L., Hruda, L.Z., Anderman, E., Anderman, L., Freeman, K. E.,…Urdan, T. (2000). *Manual for the Patterns of Adaptive Learning Scales.* Ann Arbor, MI: University of Michigan.

Mowbray, C.T., Holter, M.C., Teague, G.B., & Bybee, D. (2003). Fidelity criteria: Development, measurement, and evaluation. *American Journal of Evaluation, 24*(3), 315-340.

Pedersen S., Arslanyilmaz, A., Williams, D. (2009). Teachers' assessment-related local adaptations of a problem-based learning module. *Education Technology Research and Development, 57*, 229-2409.

O'Donnell, C.L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K-12 curriculum intervention research. *Review of Educational Research, 78*(1), 33-84. DOI: 10.3102/0034654307313793

Wiggins, G., & McTighe, J. (2005). *Understanding by design* (2nd ed.). Alexandria, VA: ASCD.