Revealing effects of changes in middle school science teachers’ practices

Gordon J. Aubrecht, II1, Jennifer Esswein2, Jessica Creamer3, and Bill Schmitt4
1Department of Physics, Ohio State University, Marion, OH 43302 and Columbus, OH 43210
2Internal Evaluator, Portland, OR 97204
3Education Specialist, Upper Sandusky, OH 43351
4Science Center of Inquiry, 13225 N. Verde River Dr. Suite 104, Fountain Hills, AZ 85268

Abstract. Grant agencies are requiring documentation that goes beyond anecdote. We work with inservice middle-school and high-school teachers in two high-needs urban school districts in Ohio. Teachers who attend the summer institutes received at least 133 hours of professional development, which leads to changes in classrooms. We report on several ways we measure changes in teacher practice, including self reports and staff observations.

PACS: 01.40.-d, 01.40.J-, 01.40.gb, 01.40.jh

I. INTRODUCTION

The Inquiry Model for Professional Action and Content-rich Teaching (IMPACT) was conceived by Principal Investigator (PI) Aubrecht and Science Partner (SP) Schmitt in 2008 as a professional development (PD) initiative to reform the teaching practices of middle- and high school science teachers. [1] The program was first implemented in Marion City Schools, an urban school district located in the Rust Belt of north-central Ohio, classified as “high-needs” by the Ohio Department of Education (ODE). After documented successes in increased student standardized test results and gains in increased teacher content knowledge, the program was expanded to the South-Western City Schools (SWCS) in 2014 in an attempt to replicate IMPACT’s successes in another ODE-classified “high needs” school system.

IMPACT’s primary goal is to promote increased student achievement by providing students the opportunity to benefit from better-prepared teachers – both pedagogically and in terms of increased teacher content knowledge. IMPACT’s triadic approach to PD involves three distinct components that, while implemented in unison, are distinct in their objectives: increased teacher-content knowledge support; the use of common formative assessments (CFAs) with students; and lesson development with an emphasis on research-based inquiry pedagogical approaches. [2] This paper does not further discuss increased teacher content knowledge or the use of CFAs in the classroom but rather focuses on reported and observable measures of teachers’ pedagogical changes.

IMPACT’s PD program for increased teacher pedagogical knowledge culminating with increased student achievement is a cyclic process that occurs throughout the calendar year. We estimate teachers attending all PD receive more than 133 hours of research-based pedagogical exposure annually: ~ 105 hours of Summer Institutes where participating teachers experience inquiry-based leaning in the role of the student in which teachers are exposed to hands-on, minds-on approaches to science instruction that honor student thinking and reasoning; ~ 63 hours of grade level meetings held monthly during the academic year where teachers work with PI and SP to collaborate on creating comprehensive multi-week content units focusing on student-centered inquiry-based instruction where the primary role of the teacher resides as the facilitator of student learning; ~ 10 hours of inclass visits from IMPACT staff which serves (among other things) to help provide feedback to specific teachers’ instructional approaches and allows the opportunity for staff to provide support for teachers implementing the inquiry-based content lessons they created into the classroom environment. The IMPACT team also uses this time to complete teacher evaluations as a measure of teacher growth.

In order to quantify changes revealed in IMPACTed teachers’ pedagogical approaches in the classroom, we rely on various diagnostics completed by the teachers throughout the 2013-15 school years, including: a curriculum and instruction diagnostic (CID), self-reports using categories from the Reformed Teacher Observation Protocol (RTOP), and staff classroom observations. [3,4] This paper discusses the results of these diagnostics.

II. RESULTS FROM THE CID SURVEY

The Curriculum and Instruction Diagnostic (CID) is a self-reported measure in which teachers quantify classroom activities and characteristics, the extent to which they use specified standards or curricula, and also report on self-efficacy in teaching science.

The following analysis provides the difference between the pre and post testing within each section and also describes how it relates to the inquiry-based professional development. The diagnostic survey questions completed by the teachers and discussed here appear in the Appendix.

A. Instructional influences

The purpose of this portion of the CID is to help portray the classroom, school, and district demographics of

39
treatment teachers. Teachers are asked to rate the effects of various instructional influences when planning for and implementing classroom instruction. Participating teachers report similar patterns in instructional influences at the beginning and end of the school year. In both fall and spring, textbooks influenced teaching the least, while district and state standards (as well as preparation of students for the next grade level) are top influencers. Interestingly, district and state test results draw the most negative influence on teaching in both cases.

B. Class instructional time

This section consists of six items, and respondents reported the amount of time their students spent in class on a given activity. They select None (0) through Considerable (50% or more time). For analysis, teachers are given a score of one through five corresponding to the task in question. Teachers are given high scores when the reported activity is more student-centered or inquiry-based than teacher-centered, and low scores when opposite. With this scoring, the highest possible score for this section is 30, and the lowest possible score is six. When the pre and posttests are compared for all respondents, the average score is nearly identical between the groups. For those who took both the pre and the posttest, however, there is a statistically significant increase from fall to spring: they reported that they spent more time on student-centered or inquiry-based activities after the year spent in the IMPACT program.

Table 1 shows the paired t-test between pre and post for those who completed both tests.

<table>
<thead>
<tr>
<th>Number of Observations</th>
<th>Average Score</th>
<th>Standard Error</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretest</td>
<td>23</td>
<td>21.22</td>
<td>0.52</td>
</tr>
<tr>
<td>Posttest</td>
<td>23</td>
<td>22.39</td>
<td>0.50</td>
</tr>
<tr>
<td>Difference</td>
<td>23</td>
<td>1.17</td>
<td>0.52</td>
</tr>
</tbody>
</table>

$t_{(df = 22)} = 2.25, p = 0.035$

C. Individual class time

This type of instruction, along with group instruction, was among the two larger changes in instruction reported.

<table>
<thead>
<tr>
<th>Number of Observations</th>
<th>Average Score</th>
<th>Standard Error</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretest</td>
<td>23</td>
<td>12.00</td>
<td>0.41</td>
</tr>
<tr>
<td>Posttest</td>
<td>23</td>
<td>14.00</td>
<td>0.51</td>
</tr>
<tr>
<td>Difference</td>
<td>23</td>
<td>2.00</td>
<td>0.50</td>
</tr>
</tbody>
</table>

$t_{(df = 22)} = 4.04, p < 0.001$

III. RESULTS FROM MODIFIED SELF-REPORTED RTOP

RTOP observation and evaluations offer the opportunity for an objective observer to quantify the extent to which a teacher employs constructivist, inquiry-based pedagogical approaches in the classroom. [4] In the self-report RTOP used by IMPACT, teachers self-assess their teaching practices using a selection of questions taken directly from the RTOP categories (17 of 25 items).

We examined what teachers reported about their practices. The maximum possible for the self-reported RTOP score was 68. Selected questions from the modified self-report RTOP are found in the Appendix. Teachers seem to think they are “reformed” (Table 5; RTOP scores over 50% are considered reformed). [4] Well-taught traditional classes might have an RTOP score of 30-40.
In order to compare teachers’ self-reported values with the traditional R TOP, the PI and trained R TOP consultant conducted unannounced classroom visits to treatment teachers and completed the full R TOP evaluations. As IMPACT is attempting to measure changes over time as a result of teachers’ participation in our PD program, we do not do specific interventions based on R TOP scores, and the individual R TOP results were not directly discussed with the observed teacher.

The PI’s original R TOP scores (all items) averaged 38 (N = 22). Scores in general increased over the project to an average of 49 (N = 31), with an effect size of 0.95. Supporting this significant change, the PI identified just 2 teachers as “reformed” in his first assessment (March 2014), while he found 16 “reformed” teachers in the latest assessment (May 2015). A consultant to the project training the PI found similar results. [5] This indicates that (1) treatment teachers often report themselves scoring higher on R TOP than the scores reported by evaluators and (2) that significant changes are occurring in teachers’ practices.

### IV. RESULTS FROM STAFF OBSERVATION

Both Aubrecht and Schmitt conducted over 100 teacher classroom observations of IMPACT teachers in both high-needs districts throughout the 2013-15 school years. These observation reports are kept simple in order for the focus to remain on coaching the participating teacher. The observation has three components: teacher approach, student engagement, and an indicator of whether or not the students exhibited reasoning skills. During the 2014-2015 school year, staff observation time was also used to perform R TOP evaluations on treatment teachers.

Using the staff observation report, the project’s internal evaluator is able to score the results of the classroom observations as a means of quantifying perceived changes in teacher practices. The rubric for classroom observations is included in the Appendix.

Points are awarded for student-centered approaches where there is evidence of high levels of student engagement in thinking and problem solving. In each of the three areas, improvement is found from the first half of the school year. Teachers show a shift from teacher-centered approaches (such as lecturing) to student-centered approaches (such as group projects/experiments), and also more evidence of student engagement. A more detailed explanation of the observation process and most recent data can be found elsewhere. [2]

### V. DISCUSSION

We present here preliminary evidence that IMPACT’s approach supports teachers changing teaching practices. Results from staff observations and R TOP scoring indicate IMPACTed teachers spend more time allowing students to work in a collaborative group setting with a focus on inquiry-based instructional pedagogical practices. It appears, however, that teachers’ self-reported R TOP scorings are higher than the findings of dispassionate observers. IMPACTed teachers appear to report themselves more “reformed” in their classroom teaching practices than they actually are. Because the CID is another self-reported measure, it is possible these numbers are inflated as well.

During the 2014-2015 school year, IMPACT was able to move from a pre-post method of data collection to a quasi-experimental design based on the fact that we were able to obtain a control group for our research. Our control group is comprised of SWCS middle school teachers who face similar demographic constraints as our treatment teachers. While we were able to obtain some preliminary data on student achievement during the ’14-’15 school year, the 2015-2016 school year will be the first year we will have teacher-level data on content knowledge and pedagogical practices from a control group. We expect more definitive evidence of changes in treatment teachers’ teaching practices to emerge in the next year as treatment teachers are continually exposed to more in-depth science content knowledge, more rigorous hands-on minds-on research-based pedagogical approaches to content and opportunities to implement reflective teaching practices.

### ACKNOWLEDGEMENTS

IMPACT has been supported since 2008 by grants from ODE using funds from the Math and Science Partnership Program

### References


Appendix

Science Curriculum and Instruction Diagnostic

(sections addressed)
Response Scale: 0 – None; 1 - Little (10% or less); 2 - Some (11-25%); 3 - Moderate (26-50%); 4 - Considerable (50% or more)

Class Instructional Time
How much of the total science instructional time do students…
6. watch the teacher demonstrate how to do a procedure or solve a problem?
7. take notes from lectures or the textbook? 0 1 2 3 4
8. complete computational exercises or procedures from a textbook or a worksheet?
9. use manipulatives, measurement instruments, and data collection devices?
10. work in pairs or small groups on science exercises, problems, investigations, or tasks?
11. reflect on their work?

In-Class Individual Student Work
When students work individually on science exercises, problems, investigations, or tasks,
how much time do they…
12. explain their reasoning or thinking in solving a problem?
13. make estimates and/or predictions?
14. work on a problem that takes at least 45 minutes to solve?
15. solve problems from a textbook or worksheet?

In-Class Pairs/Group Student Work
When students work in pairs or in small groups on science exercises, problems, investigations, or tasks, how much time do they…
16. talk about their reasoning or thinking while solving a problem?
17. make estimates and/or predictions?
18. analyze data to make inferences or draw conclusions?
19. work on a problem that takes at least 45 minutes to solve?
20. collect their own data?
21. present their information to others using computers, whiteboards, projectors, etc.?

Efficacy and Beliefs
Please circle the response which best describes you and your beliefs about science teaching. (strongly agree to strongly disagree)
33. When a student does better than usual in science, it is often because the teacher exerted extra effort.
34. Even if I try very hard, I can’t teach science as well as most other subjects.
35. When the science grades improve, it is often due to their teacher having found a more effective approach.
36. I know how to teach science concepts effectively.
37. If students are underachieving in science, it is most likely due to ineffective teaching.
38. The inadequacy of a student’s science background can be overcome by good teaching.
39. I understand science concepts well enough to be effective in teaching the concepts addressed at my grade level.1 2 3 4 5
40. The teacher is generally responsible for the achievement of students in science.

Modified self-reported RTOP
1. The instructional strategies and activities respect students’ prior knowledge and the preconceptions inherent therein.
2. Lessons are designed to engage students as members of a learning community.
3. In lessons, student exploration precedes formal presentation.
4. Lessons encourage students to seek and value alternative modes of investigation or of problem solving.
5. The focus and direction of lessons are often determined by ideas originating with students.
7. Elements of abstraction (i.e., symbolic representations, theory building) are encouraged when it is important to do so.
8. Connections with other content disciplines and/or real world phenomena are explored and valued.
9. Students make predictions, estimations and/or hypotheses and devise means for testing them.
10. Students are actively engaged in thought-provoking activity that often involve the critical assessment of procedures.
11. Students are reflective about their learning.
12. Intellectual rigor, constructive criticism, and the challenging of ideas are valued.
13. Students are involved in the communication of their ideas to others using a variety of means and media.
14. The teacher’s questions trigger divergent modes of thinking.
15. Student questions and comments often determine the focus and direction of classroom discourse.
16. Students are encouraged to generate conjectures, alternative solution strategies, and ways of interpreting evidence.
17. The teacher acts as a resource person, working to support and enhance student investigations.

Classroom Observation Rubric
1. (NA) Lesson description: NA
2. (5 pts) Describe pedagogical approach (ex. was lesson more lecture-based or more student-based?) Full points if lesson was entirely student-based. Some characteristics of this are as follows:
   i. Students working in small groups or pairs on a given problem to solve
   ii. Only ‘lecturing’ teacher is doing is to describe task/provide instructions to students
   iii. Teacher provides students with questions leading students to investigate their own questions/solutions
   Half points if class is partially teacher-led (with lecture) and half student-based. Zero points if teacher lectures for entire duration of observation
3. (2 pts) Were students engaged in lesson? Full two points for “yes.” Zero points for “no.” One point if there is some evidence students are engaged, but not fully
4. (5 pts) Did students make predictions/estimations/devise means for testing them? (yes/no with added notes if applicable) Full points if students made multiple predictions/tested predictions. Half points if students made and tested only one prediction. Zero points if there was no evidence of making and testing predictions.