Measuring Changes in High School Science Teacher Practice: Results of a Materials Science-focused Professional Development Program

Dr. Alison K. Polasik, The Ohio State University

Alison K Polasik received a B.S.E. degree in Materials Science and Engineering from Arizona State University in 2002, and M.S. and Ph.D. degrees from The Ohio State University in 2005 and 2014, respectively. She is an assistant professor of practice in the Department of Materials Science and Engineering at OSU. Dr. Polasik’s research interests include modeling of microstructure-property relationships in metals, assessment of educational outcomes, and engineering-specific epistemology in undergraduate students. Dr. Polasik is a member of ASM, TMS, and ASEE.

Alexandria Julius, The Ohio State University
Work in Progress:  
Measuring Changes in High School Science Teacher Practice:  
Results of a Materials-Science Focused Professional Development Program

Abstract

Since its inception in 2012, a materials science-focused professional development (PD) program for high school teachers has demonstrated improvements in teacher content knowledge and increased use of materials science demos and hands-on activities in the classroom. Program coordinators and researchers also noticed a change in teachers’ instructional practices, which was at first documented anecdotally. In an effort to measure these changes, the research team began using the Survey of Enacted Curriculum (SEC) and Reformed Teacher Observation Protocol (RTOP) in Year 4. A journal activity was also administered to record teacher and student activities during lessons throughout the academic year. The data from the first half of 2016-2017 academic year from these three tools was analyzed, and no clear correlation is shown between the tools on aspects of teacher practice that are measured by all three tools. The reasons for this are explored, and additional measurements and analyses are proposed to address the reasons for this.

Keywords

Materials Science, High School, Teacher Professional Development

Introduction

*Professional Development for Materials Science-Focused High School Courses* is a professional development program for high school teachers of physical science, chemistry, physics, and materials science elective courses. This program, funded by the Ohio Department of Education’s Mathematics and Science Partnerships program, was developed by faculty at The Ohio State University’s College of Engineering, Department of Materials Science Engineering. The program began in 2012 and, with some modification, has continued into 2017. The initial goals of the program were to increase teachers’ materials science content knowledge and improve student achievement. Data collected during this time demonstrated a statistically significant increase in teacher content knowledge and an increase in their use of guided inquiry and active learning activities (Polasik, 2016). Evidence has shown that guided inquiry and active teaching methods are correlated to increases in students’ content knowledge and capacity for scientific thinking (Shouse et al., 2007; 2010).

In the 2012 – 2015 academic years, the program monitored teacher use of materials science hands-on activities and their effectiveness as one metric of the degree to which the PD was impacting the classroom. As Figure 1, (Polasik, Daehn, and McCombs 2016) illustrates, the number of materials science activities increased substantially. This increase was also seen for laboratory activities deemed “scary” or dangerous by teachers because they required the use equipment the teachers feared or were unfamiliar with, e.g., blowtorch, acid, etc. Anecdotal reports by the teachers credit the program with changing their teaching practice. These results were not uniformly distributed among the teacher participants. Data collected during this period was insufficient to demonstrate that the program led to the perceived changes in teacher practice, and it did not indicate what changes the teachers made. At this point in the program the primary focus of the data collection was teacher and student learning. Measuring changes in teacher
instructional practices was of secondary interest. The overwhelming anecdotal data suggesting that the program was resulting in measurable changes in teacher practice led the research team to reassess methods to measure changes in teachers’ instructional practices.

**ACTIVITIES BY PROGRAM YEAR**

<table>
<thead>
<tr>
<th>Year</th>
<th>Total Number of Activities</th>
<th>Rated Good &amp; Very Good</th>
<th>&quot;Scary&quot; Activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year 1 ('12 - '13)</td>
<td>110</td>
<td>95</td>
<td>17</td>
</tr>
<tr>
<td>Year 2 ('13 - '14)</td>
<td>212</td>
<td>183</td>
<td>46</td>
</tr>
<tr>
<td>Year 3 ('14 - '15)</td>
<td>371</td>
<td>340</td>
<td>92</td>
</tr>
</tbody>
</table>

*Figure 1: Self-Reported In-Class activities used by teachers in each program year, by category. (Polasik et al., 2016)*

Factors that lead to effective PD, PD that leads to changes in teachers’ instructional practices, are well documented. Supovitz and Turner (2000) named the following five aspects of PD as being critical to changes in practice: immersion in inquiry experiences, questioning, and experiments—effectively modeling good instruction; intensive and sustained PD; engagement with teaching tasks that are based on experiences with students; a focus on subject-matter knowledge to increase teachers’ content knowledge; and it must be grounded in PD standards. They further recommended at least eighty hours of sustained PD. Most gains in teacher practice occur after the first 80 – 100 hours of PD (Banilower, Heck, and Weiss 2007). Additionally, the PD should connect to the curriculum (Garet 2002).

*Professional Development for Materials Science-Focused High School Courses* purposefully addressed each of these considerations. Teachers engage in 120 hours of PD each year of the program. These are distributed across an intensive materials science summer camp (40 hours), an online “Materials Science for Teachers” course offered through The Ohio State University and taught by Dr. Polasik (~45 hours), four face-to-face sessions at a local career technical school (24 hours), and classroom mentoring (8 – 10 hours depending on teacher need), and Saturday “demonstration” sessions led by materials science master teachers (~12 hours). Through the camp teachers are immersed in materials science in inquiry experiences, experiments, and materials science content. Throughout the program explicit curriculum connections are made so that teachers can connect materials science concepts to the foundational science content they teach. Teachers are given opportunities to share their classroom practices, supporting one another as they discuss what works, how they can tweak experiments and demonstrations to aid in student understanding, and where they can access needed materials. Finally, the PD program is guided and aligned to the Ohio Standards for Professional Development (Ohio Department of Education 2015).
Measuring teacher practice in science classrooms has proven to be difficult. Self-report surveys are cost effective, easy to administer, and allow for the collection of large quantities of data. However, the reliability and validity of self-report surveys is suspect (Mayer 1999). While teachers may be uniquely qualified to report on their instructional practices (Goe, Bell, and Little 2008), their self-reports may reflect their instructional intentions and perceptions rather than what is actually occurring (Copur-Gencturk, Hug, and Lubienski 2013). Classroom observations, on the other hand, are generally more objective, but are expensive and time intensive. Several research-based observation protocols have been developed in recent years: Reformed Teaching Observation Protocol (RTOP), Extended Inquiry Observation Rubric (EIQR), Science Teacher Inquiry Rubric (STIR), and the Practices of Science Observation Protocol (P-SOP) Forbes, Biggers, and Zangori 2013.)

For our purposes, we define teacher practices as the behaviors teachers engage in to plan, deliver, and reflect on their teaching. Improvement in teacher practices is being defined by changes in the frequency and nature of the teachers’ use of guided-inquiry and active learning activities.

After researching various instruments available to measure teacher instructional practices, the research team decided to use both self-report and classroom observation instruments. In an effort to measure changes in teacher practice, the team ultimately chose to use three instruments: RTOP, SEC, and an online journaling activity designed by the project team.

The RTOP was designed to quantitatively measure the extent to which classroom instructional practices align with reform principles. Reformed teaching shifts from traditional teacher-centered teaching practices to constructivist student-centered practices. The RTOP includes twenty-five items distributed across three scales: Lesson Design and Implementation (five items), Content (ten items across two subcategories), and Classroom Culture (ten items across two subcategories). Key to Lesson Design is honoring students’ prior knowledge and engaging students in experiences before expecting expertise. Content is divided into Propositional Knowledge, coherent conceptual understanding, and Procedural Knowledge, which can be described as behaviors associated with scientific habits of mind and metacognition. Classroom Culture is divided into Communicative Interactions and Student Teacher Relationships (Sawada et al., 2002).

The SEC measures the actual subject content taught and teacher instructional practices. The instrument is a self-report survey administered online. The content taught is measured in two dimensions, the topics taught and the expectations for learning. When analyzed, the data collected in these two dimensions can be used to determine the depth, breadth, and rigor of the content taught. The instructional practices measured include classroom activities, assessment, use of technology and equipment, and influences on curriculum (e.g., standards, textbooks, state or local tests) (Blank, et. al., 2001). Data from the following classroom activities sub sections are of particular interest to this study: Instructional Activities in Science, Laboratory activities, investigations, or experiments, Working in pairs or small groups, Collecting Science data or information as these best align with identified characteristics of reformed teaching measured by the RTOP.

The online journal tracks the number and type of classroom activities implemented by teachers throughout the year. Teachers fill out the survey three to four times monthly, for the same course each week (must be the same course used for the RTOP and SEC), and on roughly the same day of the week. Teachers respond to a number of straightforward questions about the types of activities included in the lesson being reported on, the teacher’s behavior, and the students’ behavior. By simply asking whether or not something occurred, it removes some of the teachers’ subjective opinions about the quality of the lesson. The particular questions regarding behavior were chosen because they identify critical aspects indicative of a class that is more active and inquiry-based. In an effort to encourage teacher participation,
the journal entry was designed to take less than five minutes to complete. The online tool is shared via email links sent weekly to the treatment and control group teachers. A copy of the journal is shown in the appendix.

**Methodology & Measurements**

In 2015 – 2016 the research arm of the program was strengthened to better evaluate the effectiveness of the program. The SEC and RTOP were added to gather quantitative data on teacher practices. A number of measurement errors were made, and the data showed no statistical differences in teacher practice. Yet, interviews with teachers and teachers’ evaluations of activities implemented in their classrooms indicated that the program is enabling teachers to make positive changes in their practice. Which is consistent with anecdotal data from previous years. While this evidence is anecdotal, the strength and consistency of these measurements support the researchers’ hypothesis that teacher practice is improving as a result of the PD. Thus, the research plan for the 2016-2017 included additional tools to analyze the general and specific teacher practices at a number of points across the academic year.

In the 2016 – 2017 school year, the research was expanded to include treatment and control teacher groups, additional measurement tools, and training for researchers and participants on the use of those tools. It is believed that this improved process will facilitate a stronger identification of the nature and causes of changes in teacher practice.

The questions specifically related to teacher practice that are identified as part of this study are as follows:

1. In what ways is the practice of the participant teachers changing? How do these changes, if present, vary across the course of the academic year?
2. Is the practice of most participant teachers in the professional development program improving?
3. Which factors (course, teacher beliefs and efficacy, teacher content knowledge, school and administrative support) correlate with improved teacher practice?

To address these questions, three tools are being used: The SEC, RTOP, and a journaling activity designed by the researchers. Each of these instruments measures slightly different aspects of teacher practice and the cognitive demand of the course for the students. The results of these tools will be compared to analyze the effects of the intervention (Research Questions 1 & 2). These results are also analyzed in conjunction with teacher and school information, teachers’ responses on a concept knowledge quiz, and a survey of teacher self-efficacy and beliefs. To improve reliability between the measurement tools, teachers who teach more than one course were instructed to choose one course in which they planned to enact the guided inquiry and active learning techniques presented in the PD program. They were further instructed to consider only this course when completing the SEC and journaling activities. This course is also the one visited by observers when completing the RTOP.

In the 2016 – 2017 academic year, three observers will observe a total of twenty teachers’ classes both at the beginning and (Sep – Oct) and end (April – May) of the year. A training program was instituted by the observation team lead, Dr. Shiverdecker, to ensure reliability of results between observers. The results will indicate the extent to which teachers have transformed their practices during the school year. The SEC is being given to all teachers (treatment and control groups) at the start and end of the academic year.

While pre- and post- academic year application of the SEC and RTOP will provide extensive measures of teacher practice in the classroom, they are not designed for repeated measures of teachers’ behaviors throughout the school year. The intent of the online journal is to track the number and type of classroom activities implemented by teachers throughout the year.
The researchers believe that including three separate measures of teacher behavior will make it possible to determine if the practice of the treatment teachers is affected by the PD. The use of the journaling activity in particular will evaluate the trends in teacher behavior over the course of an academic year. Analyzing the results of the journaling activity in conjunction with the SEC and RTOP results will indicate the validity of the tool itself.

Results

2015 – 2016 Academic Year

In the 2015 – 2016 school year, teachers reported on activities they had done the previous year. Results showed that there was a statistically significant increase in both the total number of activities implemented by teachers and the number of activities that had been deemed “scary” (Table 1). At pre-administration on average, teachers had implemented eight activities the prior school year but by the end of the school year the average number of activities had over doubled ($M = 20.8$). The average number of scary activities went from one to almost four. The changes in the average number of all activities and scary activities from pre- to post-administration had a moderate effect (effect size = .33 and .42, respectfully). There was a pronounced difference between the number of activities teachers planned to do that they considered “scary” because of potential safety issues and the number that were in fact done.

Table 1: Average number of Activities Implemented, Pre (August 2015) and Post (May 2016) Professional Development Administration

<table>
<thead>
<tr>
<th>Activities</th>
<th>Pre</th>
<th></th>
<th>Post</th>
<th></th>
<th>Wilcoxon Signed Ranks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td></td>
<td>M</td>
<td></td>
</tr>
<tr>
<td>All Activities</td>
<td>8.0</td>
<td>16.8</td>
<td>20.8</td>
<td>18.1</td>
<td>$z=3.37, p&lt;.05$</td>
</tr>
<tr>
<td>Scary Activities</td>
<td>1.1</td>
<td>2.5</td>
<td>3.8</td>
<td>3.4</td>
<td>$z=3.43, p&lt;.05$</td>
</tr>
</tbody>
</table>

Teachers were polled on their sense of efficacy and implementation of classroom activities. The results of these polls demonstrate that there is no statistically significant change in teachers’ efficacy over the course of the PD year (Table 2). This is interesting, especially considering the answer to the question “I know what is necessary to teach materials science concepts effectively”. Pre-Post concept inventories of the teachers’ content knowledge demonstrated improved understanding of materials science concepts, but these results indicate this knowledge increases did not significantly affect efficacy. From these analyses, it seems clear that one answer to research question 1 is that teacher practice is changing in a way that facilitates a greater number of hands-on and student activities.

Table 2: Teachers’ Sense of Efficacy Statistics, by Scale

<table>
<thead>
<tr>
<th>Scale/Item</th>
<th>Pre</th>
<th></th>
<th>Post</th>
<th></th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td></td>
<td>M</td>
<td></td>
</tr>
<tr>
<td>Instructional Strategies</td>
<td>7.8</td>
<td>0.8</td>
<td>7.4</td>
<td>0.9</td>
<td>0.16</td>
</tr>
<tr>
<td>Classroom Management</td>
<td>7.6</td>
<td>1.0</td>
<td>7.4</td>
<td>1.1</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Student Engagement</td>
<td>6.9</td>
<td>1.2</td>
<td>6.8</td>
<td>1.2</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Total Teachers’ Sense of Efficacy</td>
<td>7.4</td>
<td>0.9</td>
<td>7.3</td>
<td>0.9</td>
<td>0.02</td>
</tr>
<tr>
<td>How much can you do to get students to follow lab safety rules?</td>
<td>8.0</td>
<td>0.9</td>
<td>7.9</td>
<td>1.2</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Note: The items used in this analysis were based on a 9-point scale, where 1= nothing; 9=a great deal
Preliminary Analysis: 2016 – 2017 Academic Year

To quantify teacher practice, the data from the SEC, RTOP, and a journal activity were compared on two aspects of reformed teaching that were probed in both the SEC and journal survey tools. The analysis used to compare the SEC and Journal to the RTOP is summarized in Table 2. The instructional Activities in Science portion of the SEC asks teachers to report on the relative amount of instructional time spent by students on different types of activities in the previous instructional year (in this case, the 2015 – 2016 AY). Each of these items was measured on a scale of 0 to 4, where 0 corresponds to no time spent and 4 corresponds to considerable (>50%) time spent. Individual items in the survey were identified as corresponding to either procedural knowledge (IP.3, IP.8, IP.12, IPa.2, IPa.3, IPa.5, IPa.6, IPa.7, IPa.8, IPa.9, IPa.10) or communicative interactions (IP.4, IP.7, IP.9, IPa.3). These items were quantified by 1) average score over all (out of 4) and 2) number of items for which the score was 3 or higher was tabulated. The journal activity asked teachers to simply respond whether a given activity occurred. The number of activities that corresponded with each aspect of reformed teaching was scaled such that it corresponded to a number between 0 and 4, with 4 being a highly reformed lesson. In this analysis, the RTOP is effectively the standard for measuring teacher practice against which two methods of self-reporting – the SEC and the Journal – are compared. This provides an intermediary question related to the methodology used to address RQ 1-3, specifically: can self-reported measures of teacher practice be correlated to RTOP scores?

Table 3: Description of comparison between SEC, RTOP, and Journal instrument scores.

<table>
<thead>
<tr>
<th>Procedural Knowledge</th>
<th>Journal</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SEC</strong></td>
<td><strong>RTOP</strong></td>
</tr>
<tr>
<td>How much of science instructional time in class do students spend…</td>
<td>Evaluates the kinds of processes the student are asked by the teacher to use during the lesson. Procedural Knowledge is related to a teacher’s use and understanding of reformed pedagogy.</td>
</tr>
<tr>
<td>Use hands-on materials, reflect on their work, solve science problems that require novel or non-formulaic thinking, Analyze data to make inferences or draw conclusions, change a variable in an experiment to test a hypothesis</td>
<td>Asked questions, guided or probed student thinking, provided students with opportunities to gather data, encouraged students to generate conjectures and alternate strategies</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Communicative Interactions</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SEC</strong></td>
<td><strong>RTOP</strong></td>
</tr>
<tr>
<td>How much of science instructional time in class do students spend…</td>
<td>Evaluates the nature of communication between students and between the students and teacher “Lessons where teachers characteristically speak and students listen are not reformed”.</td>
</tr>
<tr>
<td>Present or demonstrate to others, engage in a writing process to support arguments with evidence, Work in pairs or small groups, explain their reasoning or thinking using several sentences orally or in writing</td>
<td>Teacher: Asked questions to probe or guide student thinking, encouraged students to generate conjectures and alternate strategies</td>
</tr>
<tr>
<td></td>
<td>Student: Engaged in discussions with peers, asked questions of peers and/or teacher</td>
</tr>
</tbody>
</table>
The results are shown in Table 4. This analysis indicates that there is no discernible correlation between teachers’ RTOP scores in Fall of 2016 and their self-reported classroom analysis provided by the SEC and journal activity, as shown in Figure 2. This is true regardless of how the SEC data is analyzed (average score or the number of activities for which a moderate or considerable amount of time is spent). More surprising, there is no clear correlation between the Journal activity and the SEC (Figure 3).

Table 4: Data for RTOP, Journal, and SEC scores for a subset of treatment and comparison teachers. The SEC score was quantified in two ways, 1) as an average of the relevant items, and 2) as the number of times each item was rated “moderate” or “considerable”

| Teacher | Procedural Knowledge | | | Communicative Interactions | | |
|---------|----------------------|------|------|--------------------------|------|------|------|
|         | RTOP Avg | Journal Avg | SEC Avg | SEC Mod/Con | RTOP Avg | Journal Avg | SEC Avg | SEC Mod/Con |
| A       | 1        | 2        | 1.09   | 0         | 1.8       | 2.25     | 1.13   | 1         |
| B       | 0.8      | 0.8      | 1.32   | 2         | 2.2       | 2.2      | 1.38   | 1         |
| C       | 0.4      | 2        | 1.23   | 0         | 0.8       | 2        | 0.88   | 0         |
| D       | 0.8      | 2.25     | 1.41   | 1         | 2         | 2.63     | 1.13   | 1         |
| E       | 2        | 1.25     | 1.32   | 2         | 2.8       | 1.25     | 1.5    | 1         |
| F       | 0.6      | 1.24     | 1.64   | 1         | 1.4       | 1.44     | 2.5    | 3         |
| G       | 1.4      | 3        | 0.18   | 0         | 0.6       | 3.0      | 0.25   | 0         |
| H       | 1        | --       | 1.95   | 2         | 1         | --       | 2.13   | 2         |
| I       | 0.2      | --       | 2.64   | 2         | 0.6       | --       | 2.5    | 3         |
| J       | 0.4      | 0.5      | 1.82   | 0         | 1.4       | 1.5      | 1.38   | 1         |
| K       | 0        | 2.5      | 1.23   | 1         | 0.4       | 2.5      | 0.75   | 0         |

Figure 2: Comparison between the SEC, RTOP, and Journal. The SEC and Journal scores determined from the coding procedure are compared to the RTOP scores taken in October 2016 for selected teachers. No clear correlation is seen in the scores regardless of which quantitative method is used.
In the 2016 – 2017 school year, the same tools are being used to measure teachers’ use of hands-on activities and self-efficacy. In addition, focus groups and a journal activity are being used to evaluate the nature and timing of changes in practice and to evaluate the reasons for lack of changes in overall teacher efficacy.

Discussion and Conclusion

The PD program outlined in this report was originally constructed to improve teacher and student content knowledge. Anecdotal evidence collected in years 1 – 3 (July 2012 through May 2015) indicates that significant changes in teacher practice were observed for a majority of participants. Measuring these changes using established instruments (SEC and RTOP) has been inconclusive. To address this, an additional instrument was developed and implemented in the 2016-2017 academic year. Preliminary analyses of the 3 instruments show that there is no clear correlation between the individual tools on the two aspects of classroom they all at least partly address: Procedural Knowledge of the teacher and Communicative Interactions within the classroom. There are a number of potential reasons for this, not the least of which is that each tool is fundamentally different. In this analysis, the RTOP was treated as the de facto “standard” of actual teacher practice and the self-reported SEC and Journal surveys were compared to it.

- The SEC administered in August and discussed in this report also deals with the 2015 – 2016 academic year, and the assumption that teachers’ practice did not change much from year to year could be invalid. The cognitive load of the SEC is quite high, and teachers may not be able to provide reliable data on classroom practice from a previous year.
The RTOP is administered only once for this analysis because of time and logistical constraints. Teachers were asked to ensure that the lesson observed was as “typical” as possible, but it is possible that the inherent differences between the observed lesson and the overall classroom and teacher activities could cause significant differences in the RTOP scores.

The Journal activity does not request information on the degree to which any student or teacher activity occurs during a lesson. While this increases compliance on the part of the participants, it makes it difficult to determine whether the teachers are interpreting individual activities the same way.

These findings are consistent with the work of Copur-Gencturk, Hug, and Lubienski (2013) who also found discrepancies between teacher self-reported data and data from classroom observations. One possible explanation of the differences may be that teachers report on what they intended to do rather than what was actually carried out.

To address some of the issues highlighted by this analysis, additional measurements are being added to the research plan for spring 2017. Teachers will be asked to fill out a detailed journal for the same lesson given when the RTOP observer is present. The program is also collecting information about teachers’ content knowledge, beliefs, and self-efficacy at the start and end of the academic year. The data will be examined for correlations between the measurements themselves and various factors believed to influence teacher practice (number of years teaching, courses taught, school type, etc.). While any correlations found are likely to be tenuous, these results will lead to additional research questions that can be incorporated in the 2017–2018 program year.

References


Journal Entry

Class Information:
Teacher Name ____________________
Topic (e.g., ceramics, atomic structure, etc.) ____________________
Date ________________

Type of activity:
☑ Demonstration
☑ Lab
☑ Classroom discussion
☑ Lecture
☑ Other

Description of Activity _____________________________________________
___________________________________________
___________________________________________

Teacher Behavior
☑ Lecture
☑ Asked questions to guide/redirect student thinking
☑ Asked questions to probe student thinking
☑ Provided students with data
☑ Provided students with opportunities to gather data
☑ Encouraged to generate conjectures, alternative solutions/strategies, and/or consider different ways of interpreting evidence

Student Behavior
☑ Gathering data
☑ Analyzing data
☑ Drawing conclusions
☑ Engaging in discussions with peers
☑ Made predictions, estimations, and/or hypotheses
☑ Asking questions of peers and/or teacher
☑ Used mathematics
☑ Chose which data to collect

Used ASM demo/activity from book? Which one? ________________________________
☐ Yes
☐ No

What other information should we know about this lesson? ________________________________
___________________________________________
___________________________________________