POWERFUL PEDAGOGICAL PRACTICES

EVALUATION:

ANNUAL EVALUATION REPORT – YEAR 1

July 2009
T.2009.04

Ximena Uribe-Zarain, Ph.D., Educational Researcher

Report prepared for:
Mathematics and Science Education Resource Center
Acknowledgements: The conduct of these evaluation activities would not have been possible without the cooperation of participating teachers, as well as the R&D Center staff members who coordinated and conducted the observations.
This report includes four sections: The first section briefly describes the program. The second section includes the methodology of the evaluation. The results of the spring observations are described in the third section. Finally, the conclusion is presented in the fourth section.

Introduction

This evaluation report, prepared by the Delaware Education Research and Development Center, includes a description of the performance of a group of Delaware mathematics teachers who participated in the Powerful Pedagogical Practices (P-cubed) project in the school year 2008-2009. This is the first year of the evaluation.

The P-Cubed program is sponsored and supported by the Delaware Mathematics Coalition in partnership with the University of Delaware's Mathematics and Science Education Resource Center (MSERC) and is funded by a state-awarded U.S. Department of Education Mathematics and Science Partnership grant.

P-Cubed is a program aimed to improve secondary math education by focusing on the manner in which secondary teachers approach mathematics instruction. Teachers are urged to seek the best way to improve their own teaching by focusing on the mathematical reasoning of students. The intention is that when teachers become genuinely curious about student thinking, they would then rethink their pedagogy.

The program focuses on ten "high leverage behaviors". The behaviors are strategies designed to facilitate discussion and mathematical reasoning among students. The ten high leverage behaviors are as follows:

1. Maximize opportunities for students to engage in discussions involving proof, justification, and argumentation during the lesson.
2. Increase the percentage of time spent listening and probing student thinking.
3. Interview students about their solution strategies for problem-based tasks and focus on questioning strategies.

4. Increase the use of descriptive feedback in order to promote student understanding.

5. Increase opportunities for students to share strategies with their classmates.

6. Focus on increasing the capacity to plan for and anticipate students’ responses for planned mathematical tasks.

7. Promote use of multiple approaches to solving problems.

8. Co-plan, observe, debrief, and reflect about the implementation of a problem-based lesson.

9. Promote a culture where students begin to rely more and more on each other as mathematical authorities in the classroom.

10. Collect data regarding the nature and level of reasoning required to answer questions in an effort to increase the percentage of higher order questions.

The role of the Delaware Education Research and Development Center was to evaluate teacher instruction. The evaluation questions are:

a. What is the quality of mathematics instruction provided by secondary teachers participating in the Powerful Pedagogical Practices (P-cubed) program?

b. How does instruction in P-cubed classroom compare with instruction of a comparable group of teachers who did not participate in the program?
Methodology

To assess the quality of teaching, we used the observation protocol called “Determining the
Quality of Mathematics Instruction”. This observation protocol was developed by University educators
from the Mathematics & Science Education Resource Center in conjunction with researchers from the
Delaware Education Research and Development Center. The protocol consists of the three main
components: The design and implementation of the lesson, mathematics content, and classroom culture.

The items or questions for each of the components are as follows:

The design and implementation of the lesson:
1. Teacher clearly defines and communicates a purpose of the lesson.
2. Teacher effectively engages students with important ideas.
3. Teacher provides adequate time and structure for investigation and exploration.
4. Teacher provides adequate time and structure for "wrap-up."
5. Teacher achieves a collaborative approach to learning.
6. Teacher enhances the development of student understanding.
7. Teacher assesses the students' level of understanding.
8. Teacher plans and/or adjusts instruction based on students' level of understanding.

Mathematics content:
1. The content is balanced between conceptual understanding and procedural fluency.
2. The content is challenging and accessible to the students.
3. Teacher provides content information that is accurate.
4. Elements of mathematical abstraction are included when appropriate to do so.
5. Appropriate connections are made to other mathematics and/or to real world content.

Classroom culture:
1. Active participation of ALL is expected and valued.
2. There is a climate of respect for students' ideas, questions, and contributions.
3. Teacher's classroom management style/strategies enhance productivity.
4. The classroom climate is encouraging to students
5. Intellectual rigor and/or the constructive challenge of ideas are evident.

Using the “Determining the Quality of Mathematics Instruction” protocol two observers were
trained until they achieved an adequate inter-rater reliability. During observations, items were scored
using three principal descriptors, “close to ideal,” “getting there,” and “not even close.” While these
concepts are illustrated through examples within the context of each of the separate indicators, it is
possible to characterize them in more general terms. An indicator is rated as “close to ideal” if there is a good bit of strong supporting and little or no contradictory evidence. “Getting there” suggests a convergence on exemplary practice but also an incomplete realization thereof. Practices that are clearly at odds with the ideal within an indicator may still be present but no longer represent the norm. Teaching that is rated as "not even close," however, is consistently impoverished with little indication of progress toward the exemplary.

In February and March of 2009, as part of the first year of the P-Cubed Program evaluation, observers were sent into math classrooms to gather data about math instruction across the state. We gathered data from 39 teachers; 26 teachers were part of the program and 13 were randomly selected as control teachers. The control teachers were selected from the same schools as the teachers in the program.

The lessons observed occurred in seventh to eleventh grade classrooms, and ranged from 45 to 105 minutes in length (average = 69 minutes). There were between seven and 27 students in each classroom with an average of about 9 students. During the time of the observations, teachers were targeting anywhere from one to four high leverage instructional behaviors (see above). Sixty-one percent of them were working only on one, 17% were working on two, 11% were working on three, and 11% were working on four behaviors at the time. All together, teachers chose to focus on only five out of the ten possible behaviors. Almost half of the teachers chose to increase opportunities for students to share strategies with their classmates (high leverage behavior 5), a third chose to promote a culture where students begin to rely more and more on each other as mathematical authorities in the classroom (high leverage behavior 9), and three behaviors were equally common: a sixth of the teachers chose to promote use of multiple approaches to solving problems (high leverage behavior 7), a sixth chose to work on maximizing opportunities for students to engage in discussions involving proof, justification, and
argumentation during the lesson (high leverage behavior 1), and also a sixth decided to effectively engages students with important ideas (high leverage behavior 2)\(^1\).

**Results**

The results of the teacher observations are presented in this section. Percentages of teachers rated in each category: “close to ideal,” “getting there,” and “not even close” as well as instances where teachers were rated in the middle of the categories (e.g. in between “close to ideal” and “getting there”) are represented in graphs throughout this section. The graphs include P-cubed and control teachers’ observation scores side by side for each question of the three components. In addition to the graphs of frequencies, we included the analysis of statistical significance for the differences between the two groups for each question of the three components. We used a scale from one to five, with the highest score of five assigned to “close to ideal.” Finally, we focused on the behaviors teachers were working on and their relationship to their performance in certain questions of the protocol.

*The design and implementation of the lesson*

The first domain, pertaining to the design and implementation of the lesson, is represented in Figure 1. Overall, the frequency of a “close to ideal” performance was higher for P-cubed teachers than for control teachers. Similarly, the frequency of a “not even close” performance was lower for P-cubed teachers than for control teachers. In general, between 60 and 90% of P-cubed teachers were classified as “close to ideal” or getting there” for all the items in this section. Up to two thirds of control teachers were classified as “not even close” in this section. More specifically, 39% of P-cubed teachers were classified as “close to ideal” in assessing the students’ level of understanding; including prior experiences, preparedness and learning styles (Question 6), while only 15% of control teachers did. At the same time, 39% of control teachers were classified as “not even close” in this same item.

---

\(^1\) Percentages do not add up to 100 because some teachers selected more than one high leverage behavior at once.
Are the differences between P-cubed and control teachers' performance statistically significant?

For the design and implementation of the lesson section, P-cubed teachers scored higher than their control peers. However, only one indicator showed statistical significant difference between the two groups. P-cubed teachers \((m= 3.577, sd=1.301)\) assessed the students' level of understanding more often than control teachers \((m= 2.385, sd=1.502)\). This difference was statistically significant \((t=2.56, \ p < .05)\). P-cubed teachers emphasized higher order questions that required students to synthesize information or form generalizations more frequently than control teachers. Teachers in the program also asked their
students questions such as “why do you think so?” more often than their counterparts. Still, P-cubed teachers’ scores could be higher; this indicator received is one with the lowest scores for P-cubed teachers in this section.

<table>
<thead>
<tr>
<th>Question</th>
<th>group</th>
<th>mean</th>
<th>std dev</th>
<th>t</th>
<th>sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>P-cubed</td>
<td>4.192</td>
<td>1.327</td>
<td>0.26</td>
<td>ns</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>4.077</td>
<td>1.188</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>P-cubed</td>
<td>4.240</td>
<td>1.165</td>
<td>0.95</td>
<td>ns</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>3.833</td>
<td>1.337</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>P-cubed</td>
<td>3.520</td>
<td>1.584</td>
<td>1.26</td>
<td>ns</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>2.833</td>
<td>1.467</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>P-cubed</td>
<td>3.462</td>
<td>1.529</td>
<td>1.06</td>
<td>ns</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>2.923</td>
<td>1.441</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>P-cubed</td>
<td>3.654</td>
<td>1.522</td>
<td>1.83</td>
<td>ns</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>2.692</td>
<td>1.601</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>P-cubed</td>
<td>3.577</td>
<td>1.301</td>
<td>2.56</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>2.385</td>
<td>1.502</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>P-cubed</td>
<td>3.885</td>
<td>1.275</td>
<td>1.55</td>
<td>ns</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>3.231</td>
<td>1.166</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>P-cubed</td>
<td>4.385</td>
<td>0.983</td>
<td>1.04</td>
<td>ns</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>4.000</td>
<td>1.291</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TABLE 1. Descriptive statistics and mean differences for design and implementation of the lesson

P-cubed teachers were rated highly for communicating the purpose of the lesson (Question 1), engaging students with important ideas (Question 2), and adjusting the lesson based on students’ understanding (Question 8). This last question speaks to the core of the program. Teachers should focus on the mathematical reasoning of students. As stated before, the intention of P-cubed is that teachers will consider student thinking, and then reconsider and adjust their teaching. This seems to be happening in P-cubed classrooms.
There is still room for improvement in many areas. Less than 50% of P-cubed teachers performed “close to ideal” in five out of the eight questions (Question 3 to 7) in the section of design and implementation of the lesson. Time management could be improved by allocating more time for investigation, exploration, and for wrapping up the lesson. Exploration and investigation are key components of the desired learning process (Question 3). And the “wrap-up” should be an opportunity for the teacher to orchestrate the presentation of student ideas that have resulted from the exploration phase (Question 4) and, in so doing, draw out the important mathematical ideas from the lesson. Teachers need to assign more time to these two very important parts of the lesson. Another area for improvement is collaborative approach of learning (Question 5). Here, we not only look for teachers encouraging students to interact with each other, but also for the interaction to be meaningful and rich. Almost 20% of P-cubed teachers were classified as “not even close.” Finally, we have students’ understanding. We know that P-cubed teachers adjust their lessons according to their students’ understanding, but they could enhance students understanding (Question 6) better by asking higher order questions more often and addressing misconceptions promptly. Also, teachers could ask more question to expand students understanding (Question 7) and to make a better judgment of students’ understanding.

Mathematics content

Items regarding mathematics content of the lesson are represented in Figure 2. In four out of the five indicators which comprise this section, P-cubed teachers had a higher percentage of performing “close to ideal” than their peers in the control group. More than three quarters of the P-cubed teachers were classified “close to ideal” or “getting there” for all the questions. There are also more control teachers performing at the lowest level in all five questions. More specifically, more than 50% of P-cubed teachers scored at the highest levels on working with elements of mathematical abstraction (Question 4);
while less than 25% of control teachers did. P-cubed teachers were more successful moving their
students from abstraction to prediction or application and students were able to apply the idea in a new
context.

![Graph showing comparison of P-cubed and control group](image)

**FIGURE 2. Mathematics content**

There were no statistically significant differences between teachers in the program and the control
group. Means, standard deviations and the results from the comparison between the two groups are
presented in Table 2.
Regarding P-cubed teachers frequencies alone, more than 50% of the teachers were classified as “close to ideal” in all the questions in this category. And very low percentage of them was classified as “not even close.” Nevertheless, there is room for improvement especially when it comes to teachers presenting content that is challenging and accessible at the same time (Question 2) and helping students to move from concrete to abstraction (Question 4).

<table>
<thead>
<tr>
<th>Question</th>
<th>group</th>
<th>mean</th>
<th>std dev</th>
<th>t</th>
<th>sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>P-cubed</td>
<td>4.077</td>
<td>1.354</td>
<td>0.93</td>
<td>ns</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>3.615</td>
<td>1.660</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>P-cubed</td>
<td>4.000</td>
<td>1.265</td>
<td>0.85</td>
<td>ns</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>3.615</td>
<td>1.446</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>P-cubed</td>
<td>4.385</td>
<td>1.061</td>
<td>0.00</td>
<td>ns</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>4.385</td>
<td>0.768</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>P-cubed</td>
<td>3.833</td>
<td>1.494</td>
<td>1.79</td>
<td>ns</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>2.923</td>
<td>1.441</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>P-cubed</td>
<td>4.360</td>
<td>1.186</td>
<td>-0.54</td>
<td>ns</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>4.583</td>
<td>1.165</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TABLE 2. Descriptive statistics and mean differences for mathematics content

Classroom culture

The third section Classroom Culture is represented in Figure 3. The first three indicators were rated very highly for both groups. The vast majority of teachers in both groups were classified as “close to ideal” and a very low percentage was classified as “not even close.” These three questions deal with student participation, climate of respect in the classroom, and classroom management. It is apparent that active participation of all students is almost always expected or valued (Question 1). Also, most lessons showed a climate of respect for students’ ideas, questions and contributions with very few rated “not even close” (Question 2). Almost 90% of P-cubed teachers were classified as “close to ideal” in regards to classroom management style (Question 3).
The last two questions are different. Question 4 is about teacher acting as facilitator so that students can play the role of problem solvers. Question 5 is about intellectual rigor and teachers expecting students to explain their reasoning. Almost half of the teachers in the program were classified as "close to ideal" where less than 10% of control teachers did in both instances.

![Bar chart showing comparison between P-cubed and control group for classroom culture questions](image)

**FIGURE 3.** Classroom culture: P-cubed and control group

In the last two questions, P-cubed teachers scored better than their peers, both with higher percentage of "close to ideal" and a lower percentage of "not even close." These differences are also statistically significant ($t=2.45, \ p < .05$ for Question 4 and $t=3.61, \ p < .05$ for Question 5). P-cubed teachers showed a better performance when acting as facilitators in problem solving ($m=3.885, \ sd=1.211$).
than the control teachers ($m=2.923$, $sd=1.038$). The greatest difference was observed in matters of intellectual rigor or constructive challenge of ideas (Question 5). The frequency in which P-cubed ($m=3.731$, $sd=1.282$) classrooms students' conjectures were explored and students were held to the standard of justification and proof was significantly higher than classrooms of control teachers ($m=2.231$, $sd=1.092$).

<table>
<thead>
<tr>
<th>Question</th>
<th>group</th>
<th>mean</th>
<th>std dev</th>
<th>t</th>
<th>sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>P-cubed</td>
<td>4.538</td>
<td>0.706</td>
<td>0.86</td>
<td>ns</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>4.308</td>
<td>0.947</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>P-cubed</td>
<td>4.538</td>
<td>1.029</td>
<td>-0.12</td>
<td>ns</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>4.583</td>
<td>1.165</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>P-cubed</td>
<td>4.769</td>
<td>0.652</td>
<td>0.65</td>
<td>ns</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>4.615</td>
<td>0.768</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>P-cubed</td>
<td>3.885</td>
<td>1.211</td>
<td>2.45</td>
<td>&lt;.05</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>2.923</td>
<td>1.038</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>P-cubed</td>
<td>3.731</td>
<td>1.282</td>
<td>3.61</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>control</td>
<td>2.231</td>
<td>1.092</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### TABLE 3. Descriptive statistics and mean differences for classroom culture

Still, the performance on these two last questions could be improved. P-cubed teachers did a-rate higher than control teachers. But still in a third of P-cubed lesson we observed, problem solving was teacher-directed or highly scaffolded. Many teachers outlined step by step solutions in problems they posed to students. The same situation happened with intellectual rigor, in a third of the lessons we observed, students were expected to explain their reasoning but they were not held to the standard of justification and proof.

*High leverage behaviors and performance*
We matched each high leverage behavior with indicators in each component of the observation protocol (see table 4). Based on this matching, we assessed how P-cubed teachers working on the most popular\(^2\) behaviors performed.

<table>
<thead>
<tr>
<th>High Leverage Behavior</th>
<th>Design and implementation</th>
<th>Mathematics content</th>
<th>Classroom culture</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Maximize opportunities for students to engage in discussions involving proof,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>justification, and argumentation during the lesson.</td>
<td>3, 6</td>
<td>1, 5</td>
<td></td>
</tr>
<tr>
<td>2. Increase the percentage of time spent listening and probing student thinking.</td>
<td>3, 6, 7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Interview students about their solution strategies for problem-based tasks and</td>
<td>3, 5</td>
<td>1, 2</td>
<td></td>
</tr>
<tr>
<td>focus on questioning strategies.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Increase the use of descriptive feedback in order to promote student understanding.</td>
<td>7, 8</td>
<td>2, 2</td>
<td></td>
</tr>
<tr>
<td>5. Increase opportunities for students to share strategies with their classmates.</td>
<td>3, 5</td>
<td>1, 2</td>
<td></td>
</tr>
<tr>
<td>6. Focus on increasing the capacity to plan for and anticipate students' responses</td>
<td>3, 6, 7</td>
<td>4, 5</td>
<td></td>
</tr>
<tr>
<td>for planned mathematical tasks.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Promote use of multiple approaches to solving problems.</td>
<td>4, 8</td>
<td>3, 5</td>
<td></td>
</tr>
<tr>
<td>8. Co-plan, observe, debrief, and reflect about the implementation of a problem-based</td>
<td>2, 7, 8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>lesson.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Promote a culture where students begin to rely more and more on each other as</td>
<td>5</td>
<td>2, 1, 2, 4</td>
<td></td>
</tr>
<tr>
<td>mathematical authorities in the classroom.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Collect data regarding the nature and level of reasoning required to answer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>questions in an effort to increase the percentage of higher order questions.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TABLE 4. Relationship of high leverage behaviors and the observation protocol

\(^2\) Out of the 10 possible behaviors, teachers chose to work on five. Behaviors 5 and 9 were the most popular and 1, 2, and 7 were chosen less.
High Leverage Behavior 5: 
Increase opportunities for students to share strategies with their classmates

- The teacher provides adequate time and structure for investigation and exploration.
- The teacher achieves a collaborative approach to learning.
- Active participation of ALL is expected and valued.
- There is a climate of respect for students’ ideas, questions, and contributions.

n=13

FIGURE 4. Frequency distributions for items related to behavior 5

High Leverage Behavior 9: 
Promote a culture where students begin to rely more and more on each other as mathematical authorities in the classroom

- The teacher achieves a collaborative approach to learning.
- The content is challenging and accessible to the students.
- Active participation of ALL is expected and valued.
- There is a climate of respect for students’ ideas, questions, and contributions.
- The classroom climate encourages students

n=8

FIGURE 5. Frequency distributions for items related to behavior 9
Teachers who chose to increase opportunities for students to share strategies with their classmates (behavior 5) scored highly in most of the items related to this behavior. Just as the rest of teachers in the program, these teachers achieve a climate of respect in their classrooms. And active participation of all students is valued. It is interesting to notice that about half of these teachers did an excellent job at setting up the lesson for investigation and exploration to occur. This means that the other half has not mastered the allocating enough time for this part of the lesson and some teachers were still telling the students next steps to solve problems. Also, in a fourth of classrooms where teachers chose behavior 5 students rarely or never interacted with each other regarding the lesson. Student interaction in groups could be a good way for students to share strategies with their peers.

Teachers who chose to promote a culture where students begin to rely more and more on each other as mathematical authorities in the classroom (behavior 9) also scored well. We observed very few instances where teachers were “not even close.” Once again, a climate of respect was observed in these lessons. Still, in this case there is more room for improvement than in the previous behavior. More work needs to be done to provide students with a classroom climate that encourage them to generate ideas and question as students are problem solvers and teachers become facilitators. Finally, we saw 25% of these teachers unable to achieve a collaborative approach in their classrooms.

Conclusion

What is the quality of mathematics instruction in Delaware secondary schools participating in the P-cubed program?

Regarding the design and implementation of the lesson, P-cubed teachers showed a good performance in the following items:

1. Clearly defining and communicates a purpose of the lesson.

8. Planning and/or adjusting instruction based on students' level of understanding.
Regarding the design and implementation of the lesson, P-cubed teachers could improve their performance in the following items:

3. Providing adequate time and structure for investigation and exploration.

4. Providing adequate time and structure for "wrap-up."

5. Achieving a collaborative approach to learning.

6. Enhancing the development of student understanding.

7. Assessing the students' level of understanding.

Regarding the mathematics content of the lesson, P-cubed teachers showed a good performance in the following items:

1. The content is balanced between conceptual understanding and procedural fluency.

2. The content is challenging and accessible to the students.

3. Teacher provides content information that is accurate.

5. Appropriate connections are made to other mathematics and/or to real world content.

Regarding the mathematics content of the lesson, P-cubed teachers could improve their performance in the following item:

4. Elements of mathematical abstraction are included when appropriate to do so.

Regarding the classroom culture of the lesson, P-cubed teachers showed a good performance in the following items:

1. Active participation of ALL is expected and valued.

2. There is a climate of respect for students' ideas, questions, and contributions.

3. Teacher's classroom management style/strategies enhance productivity.

Regarding the classroom culture of the lesson, P-cubed teachers could improve their performance in the following items:

4. The classroom climate is encouraging to students
5. Intellectual rigor and/or the constructive challenge of ideas are evident.

*How does instruction in P-cubed classroom compare with instruction of a comparable group of teachers not participating in the program?*

For the design and implementation of the lesson section, only one indicator showed statistically significant difference between the two groups. P-cubed teachers assessed the students’ level of understanding more often than control teachers. P-cubed teachers emphasized higher order questions that required students to synthesize information or form generalizations more frequently than control teachers. Teachers in the program also more frequently asked their students questions to explain their responses.

There were no statistically significant differences in the mathematics content of the lesson section between teachers in the program and the control group.

In the classroom culture section, P-cubed teachers scored better than their peers. The frequency in which P-cubed classrooms students’ conjectures were explored and students were held to the standard of justification and proof was significantly higher than classrooms of control teachers.

In summary, teachers in the program are assessing students' level of understanding more often than control teachers. This is important because the aim of P-cubed is that teachers will improve their pedagogy by paying attention and identifying the mathematical reasoning of students so that they can adjust their lesson. Two other areas are vital to achieve the goals of the program. The first area is having teachers acting as facilitators in problem solving. This means that teachers should not tell students the next steps to solve a problem. The second area is intellectual rigor or constructive challenge of ideas. Teachers can achieve this by routinely asking their students not only to explain their reasoning, but also to defend it through justification and proof. In these two areas P-cubed teachers performed statistically significantly better that their peers, but, given that their scores were still not high, they could continue to improve their powerful pedagogical practices.