November 2006

The Gilbert Elementary Science Program

Providing Students With High Quality Science Learning

This evaluation brief is supported by the National Science Foundation as part of its Local Systemic Change Initiative. The brief highlights the current status of science teaching in the Gilbert, Arizona School District. The brief uses data from more than 50 observations of science lessons to describe how the use of nationally recognized curriculum and science notebooks has created a standard of consistent, high quality science teaching in the District.

For more information see www.inverness-research.org.

Executive Summary

Inverness Research has studied the Gilbert Science program since the summer of 1998. This report presents data from observations of more than 50 science lessons in Gilbert classrooms over the last seven years. We describe the link between high quality lessons and the consistent use of nationally recognized curriculum and science notebooks. We also provide illustrations of the shift that has occurred in Gilbert classrooms towards high quality science lessons. Finally, we provide quantitative data that documents this shift towards a standard of high quality science teaching in Gilbert.

During our first round of classroom observations in the spring of 1999, we observed ten classrooms. In five of these classrooms, teachers were using nationally recognized curriculum. None of the lessons incorporated science notebooks and 30 percent of the lessons were judged to be high quality.[1] In our final round of observations in the spring of 2005, we observed 18 teachers and all of them were using nationally recognized curriculum. Eighty-nine percent of the lessons incorporated science notebooks and were judged to be high quality.

Our longitudinal data documents the shift towards high quality lessons in Gilbert as well – none of the lessons in our first or second observations of the longitudinal data set were judged to be high quality. In our final observations in spring 2005, 90 percent of these same teachers taught a lesson that was judged to be high quality.

High quality science lessons engage students with important science concepts and develop both their understanding of these concepts and their capacity to do science successfully. Our research suggests that, unlike students in 85 percent of districts nation-wide, most of Gilbert students are consistently experiencing high quality science lessons. This high standard of science teaching is clearly the result of the Gilbert science program.

[1] Horizon Research (2003) *Highlights Report – Looking Inside the Classroom: A Study of K-12 Mathematics and Science Education in the United States.*
http://www.horizonresearch.com/reports/2003/insidetheclassroom/highlights.php

Introduction

In the typical elementary classroom nationwide, science is not often taught, and if it is, the materials have usually been developed by the teacher and as a result, lack rigor and depth.[2] The Gilbert school district has an unusual situation in its elementary classrooms: teachers are regularly using nationally recognized science materials and notebooks to provide high quality science learning experiences for their students. Inverness Research, the independent evaluator of Gilbert's National Science Foundation-funded Local Systemic Change Initiative in science, has studied the Gilbert Science program since the summer of 1998. This report presents data from observations of science lessons in Gilbert classrooms. We report that teachers are using the district-designated science materials and science notebooks effectively in classrooms district-wide and we describe how the use of these two tools allows teachers to provide elementary students with unique and rich opportunities to learn science.

How we collect data during observations of science lessons

Over the last seven years, we observed more than 50 science lessons in Gilbert schools. Teachers to be observed were selected randomly from a data set of all of the K-5 teachers in the district. When we observed a lesson, we used a structured observation protocol[3] designed by Horizon Research to record data on indicators for four areas: lesson design, lesson implementation, science content, and classroom culture. We then used these data to give the lesson an overall "capsule" rating. Capsule ratings range from Levels 1 to 5, where the ratings are described as follows:

- Level 1 Ineffective instruction
  - 1a. "Passive learning"
  - 1b. "Activity for activity' sake"
- Level 2 Elements of effective instruction
- Level 3 Beginning stages of effective instruction (low, solid, high)
- Level 4 Accomplished, effective instruction
- Level 5 Exemplary instruction

Lessons are broadly categorized by Horizon as low in quality (1-2), medium quality (low-3, solid-3), and high quality (high-3, 4 and 5). Horizon describes these categories as follows: "Lessons judged to be low in quality are unlikely to enhance students' understanding of important science content or the ability to engage successfully in the process of science. At the other end of the scale, high quality lessons are structured and implemented in a manner which engages students with important science concepts; these lessons are very likely to enhance

[2] Horizon Research (2006) *LSC Research Update: Use of District-Designated Materials Increases Quality Rating of Classroom Lessons* at www.pdmathsci.net.

[3] For more information on the protocol and ratings you can go to the Horizon website at www.horizon-research. com)

student understanding of these concepts and to develop their capacity to do science successfully."[4]

The link between nationally recognized science curriculum and high quality lessons

The Gilbert Science program has worked to select instructional materials that are carefully designed, field-tested, research-based and aligned with national standards. These kits come from two nationally recognized providers, STC and Foss. Teachers began using the current kits from STC and Foss during the 1998-1999 school year, and by the year 2002-2003, three of these kits at each grade level were the designated materials for the Gilbert Science program. Horizon Research conducted a 10-year study of 88 Local System Change projects, funded by the National Science Foundation, of which the Gilbert Science program is one. They found that when teachers use district-designated instructional materials, like those from STC or Foss, their lessons were much more likely to be higher rated (Horizon Research, 2006). The study explained that teacher-made materials lack focus and rigor and also, when teachers are provided with high quality materials they tend to rely on them.[5]

During our first round of classroom observations in the spring of 1999, we observed ten classrooms. In five of these classrooms, teachers were using district-designated instructional materials from STC. In the other five classrooms, one teacher was using a MESA kit, and the others were using materials they had developed themselves. In the first year of the program (classroom observations conducted in spring 1999), 30 percent of the lessons were judged to be high quality (3 high to 5) and ratings ranged from 1b to 5.

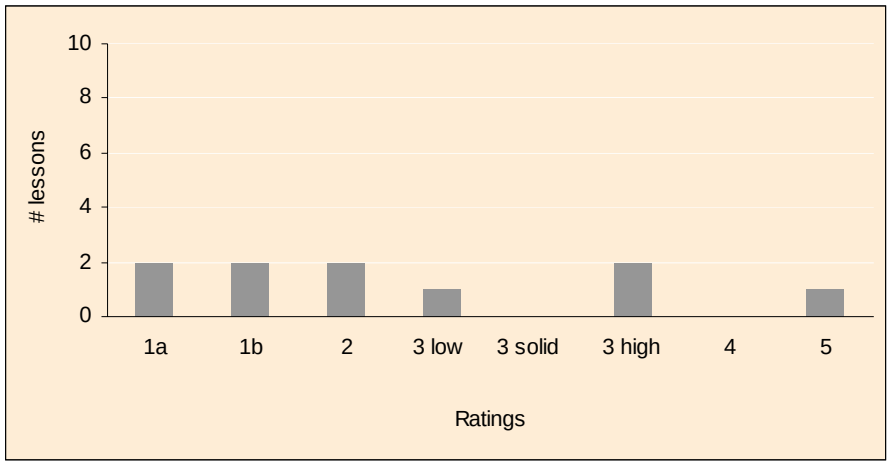
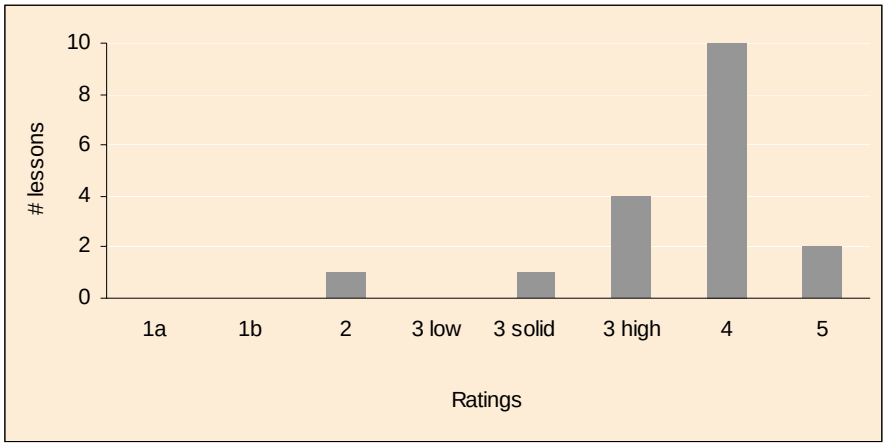


Figure 1. Random classroom observations Spring 1999 (10 lessons).

In our final round of observations in the spring of 2005, we observed 18 teachers, and all of them were using district-designated materials or supplemental materials related to the district science curriculum. Eighty-nine percent of lessons were judged to be high quality (3 high to 5) and ratings ranged from 2 to 5.

[4] Horizon Research (2003)
[5] Horizon Research (2006)

Figure 2. Random classroom observations Spring 2005 (18 lessons).



Our classroom observation data suggests that this shift is occurring district-wide. Most, if not all, Gilbert elementary teachers using the district-designated science materials, and they are using them well. The following descriptions of highly rated lessons will describe how this translates to better science instruction and more opportunities for students to learn.

The link between science notebooks and high quality lessons

This topic is covered in detail in the companion brief to this piece titled *The Gilbert Elementary Science Program: Science for Writing and Writing for Science*, which describes the use of science notebooks in Gilbert. That brief describes the results of a study of 40 science notebooks, classroom observations and interviews with teachers, and describes that when used most skillfully, science notebooks enhance student learning in both science and language arts. Writing improves students' learning of science concepts and skills, and science offers a rich and immediate context for developing writing. The evidence from this study also demonstrates that science notebooks provide a place for students to make the link from hands-on activities to science concepts, making their investigations authentic and purposeful.

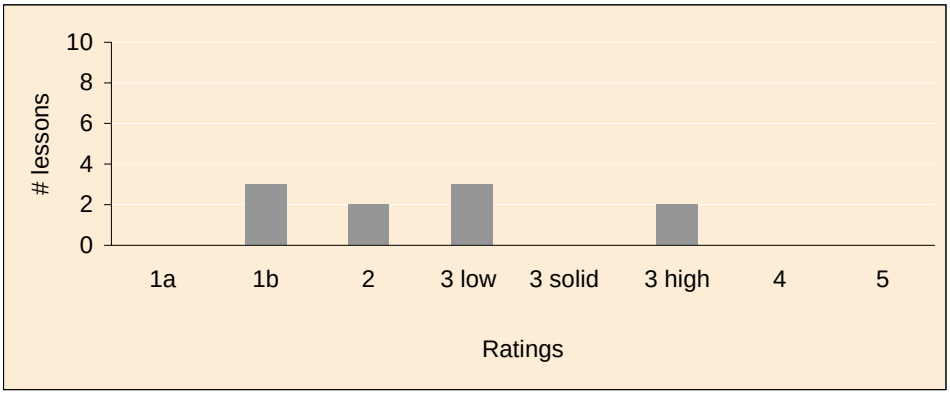
We observed ten classrooms during our first round of observations in the spring of 1999. None of the teachers used notebooks. In our final round of observations in the spring of 2005, 16 out of 18 teachers observed used science notebooks and in several classrooms the notebooks impressively enhanced the teaching and learning of science. The following vignette from an observation in the spring of 2005 illustrates how Gilbert teachers are using notebooks to help students of all abilities make meaning of their experiences during science lessons.

> *We observed a first-grade teacher teaching Lesson 10 of the Solid and Liquids unit. The students worked in pairs closely following the teacher's instructions. They investigated the properties of the liquids glue and water. Students observed these materials, touched them, and smelled them. They dipped spoons in both and the teacher said to look to see if they looked the same or different. They had a table labeled "Comparing Water and Glue" with two columns: "Different / Alike." They then did a thorough oral debriefing of what they noticed with the teacher writing what the students said on a big paper posted in front of the room. They talked about three questions: 1) How were they different? 2) What about when you felt it?; and 3) How were they alike? Then they took out their notebooks. The teacher said "Tell me what you learned from your experiment." She reminded them, "What goes on top since you're a scientist?" She asked three students to give her an example of what they might write. The teacher told them*

not to misspell words, and to use the thoughts they had generated on the big paper. Some students copied from the board; some drew very accurate pictures. A Special Education student copied earnestly and got four words down without errors. The teacher left about 20 minutes for the discussion and writing at the end which was just about right. The writing really pulled this lesson together and gave it back to the students.

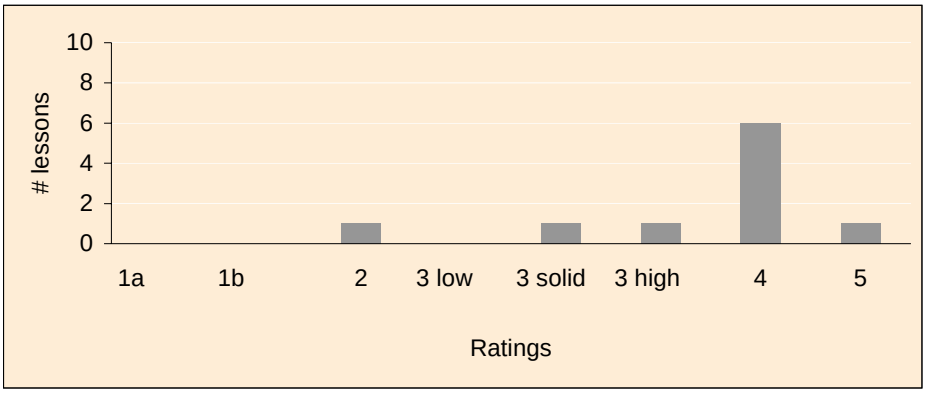
Longitudinal data suggests teachers are teaching science better as a result of the Gilbert Science program

Figure 3. Longitudinal classroom observations--first observation (10 lessons).



Over the years, we observed a subset of teachers multiple times, to collect data on whether or not participation in the Gilbert Science program had affected their science teaching. There are ten teachers for whom we have comparison data based on observations done in either 1998-99 or 1999-2000 and again in 2005.

Figure 4. Longitudinal classroom observations--final observation (10 lessons).



Nine out of ten teachers for whom we have longitudinal data showed improvement on the Horizon Research Classroom Observation Protocol over a period of four to five years. Four made marked improvement moving from capsule ratings of 1 or 2 (low quality) to ratings of 4 or 5 (high quality). Three others made notable improvements, one moved from a 1 (low quality) to a 3-high (high quality) and the other two from 3-low (medium quality) to 4 (high quality). The last teacher in the group showing

improvement had only a slight improvement from a 3-low (medium quality) to a 3-solid (medium quality).

What follows are vignettes for the first and final observations for two teachers for whom we have longitudinal data. These vignettes illustrate that, as the teachers used nationally recognized curricula and science notebooks and became more skilled in their science teaching, their lessons moved from low quality to high quality and their students had more time and opportunity for genuine exploration, reflection and learning.

<table>
<tr><td colspan="2">Teacher A</td></tr>
<tr><td>Spring 1999

Floating Egg Lesson
Capsule rating of 1

In this lesson, the students followed steps on a teacher-designed worksheet to see if a raw egg floats or sinks in salt water and regular water. The ultimate question the students were trying to answer was ‘Is it easier to swim in an ocean or a lake?’ The students filled in the blanks on the worksheet and did very little investigation or thinking on their own – the teacher set up the investigation and provided the “correct” answers.</td><td>Winter 2005

Electric Current Lesson
Capsule rating of 4

The lesson we observed is part of the Electric Currents kit from STC and is called Hidden Circuits. This lesson is designed to help the teacher assess how students apply their understanding of complete circuits. The teacher introduced the lesson via a prompting question that she had prepared for each student to glue in his/her science notebook. After a brief introduction to the hidden circuit boxes and how to record their findings, the students eagerly started their investigations. The investigative analysis of the hidden circuits was portrayed as real scientific work and the students recorded their findings in their notebooks.</td></tr>
<tr><td colspan="2">There are several key differences between these two lessons. 1) The second lesson is an integral part of a complete unit of study where students apply skills they have learned in the first half of the unit of study. In contrast, the first lesson was a stand-alone lesson. 2) In the second lesson, the students investigate in pairs and find their own answers. In the first lesson, there was no real investigation as the teacher held the answers. 3) In the second lesson, the students record their findings in their science notebooks allowing them to record the data in their own words and format. In the first lesson, they filled in the blanks in a worksheet. These key differences led to a richer lesson with many opportunities for the students to increase their knowledge about a particular scientific topic.</td></tr>
</table>

Teacher B	
Spring 1999 Energy Lesson Capsule rating of 2 The lesson we observed in Spring 1999 was one of the middle lessons from a kit, developed by MESA Arizona, about energy. Before this lesson, the students had observed "stored energy" in blown up balloons as well as learned about flashlight construction. The day we observed the classroom, the teacher lead the students through constructing a complete circuit out of a battery, battery holder, two wires, a bulb holder and a light bulb. The teacher led the students through constructing the circuits and explained the steps in such great detail that constructing the circuit became a cookbook activity with no room for exploration or learning. There was only one "right" way to make the circuit. The students then drew and labeled pictures of their circuits.	Spring 2005 Balancing and weighing Lesson Capsule rating of 5 The lesson we observed began with a review of yesterday's lesson. The lesson for today was lesson three of the Balancing and Weighing kit from STC for second grade. The object of the lesson was to use a beam, fulcrum, and 10 unifix cubes to make the beam balance with the fulcrum in the center of the beam. The students worked in pairs for about 20-25 minutes exploring the object of the lesson. They then were instructed to remember two strategies they used to balance the beam and to build one and leave it on display on their desks. During the next 25-30 minutes, the teacher led the whole class around to each of the eleven displays and they discussed them. They talked about whether the fulcrum was in the center of the beam, how the beam was balanced and the teacher used the students' suggestions to balance the displays that were not balanced. The last 20 minutes of the lesson were spent with the kids drawing their two methods of balancing their beams in their science notebook and then as a whole group listening to the teacher read a story about a gymnast balancing on one foot. The students tried balancing on one foot and the teacher helped the students make connections between them balancing and what they had learned in the lesson.
Again, there are several key differences between the lessons. 1) The first lesson illustrates how curriculum that is not from a nationally recognized source is often disjointed, with lessons jumping from one broad topic to another. As a contrast, the lessons in kits from kit developers such as FOSS and STC carefully build conceptual understanding with the first lessons laying the foundation for in-depth exploration and learning in the later lessons. 2) In the first lesson, the students are led by the teacher through specific steps to one right answer. The second lesson is a masterful example of a teacher guiding students' learning. The teacher has learned to create a culture that fosters student learning and exploration; rather than telling the students what is correct, she uses their ideas and discoveries to illustrate the central concepts of the lesson.	

Another measure of the quality is whether students remember what they learned several years later. This spring, we spoke to some middle school students, and they said the following:

> *In sixth grade, we learned how to tell time with the sun. In fifth grade, we did buoyancy experiments. I think in fourth or third grade, we did electricity. I think in first or kindergarten, we did butterflies.*

> *We didn't learn out of the textbook. We had a lot of labs. We would grow plants and light bulbs to get electricity. We wired the house. I think it was more fun, because we didn't have to learn out of the textbook and take notes. It was more hands-on.*

Conclusion

Horizon Research recently conducted a study of the nature and quality of science and math lessons nationwide.[6] In this study, Horizon staff and consultants observed 364 math and science lessons nationwide using the structured observation protocol described earlier in this report. Based on their observers' judgments, only 15 percent of math or science lessons in the United States would be considered as high quality (capsule rating of 3-high to 5).

During our first round of classroom observations in the spring of 1999, we observed ten classrooms. Thirty percent of the lessons were judged to be high quality (3-high to 5) and ratings ranged from 1b to 5. In our final round of observations in the spring of 2005, we observed 18 teachers. Eighty-nine percent of the lessons were judged to be high quality and ratings ranged from 2 to 5. Since teachers were selected randomly, this data suggests this shift is occurring districtwide. Most, if not all, Gilbert elementary teachers using the district-designated science materials and science notebooks, and they are using them to teach high quality lessons. Our longitudinal data documents the shift towards high quality lessons as well – none of the lessons in our first or second observations of the longitudinal data set were judged to be high quality. In our final observations in spring 2005, 90 percent of these same teachers taught a lesson that was judged to be high quality. As described by Horizon Research, high quality lessons engage students with important science concepts and develop both their understanding of these concepts and their capacity to do science successfully. Our research suggests that, unlike students in 85 percent of districts nationwide, most of Gilbert students are experiencing high quality science lessons. This high standard of science teaching is clearly a result of the Gilbert Science program.

[6] Horizon Research (2003)