Keeping an EYE on Things: Classroom Observation and Fidelity of Implementation

James P. Van Haneghan
University of South Alabama
jvanhane@usouthal.edu

Susan Pruet, Ph.D.
spruet@maef.net
Melissa Dean, B.S.
Mdean@maef.net
Mobile Area Education Foundation

Rhonda Neal Waltman, Ed.D.
Neal-Waltman & Associates Inc.
waltmaninc@bellsouth.net

Abstract
Introduction
Engineering curricula present a complex mix of instructional activities that share elements of traditional instruction, but also involve instructional activities that are unique. Capturing those elements in the classroom and determining their fidelity involves viewing classrooms from multiple perspectives. In our work in developing and testing Engineering modules middle school students, we have spent a great deal of time exploring how best to determine whether the intended activities in the classroom correspond with the vision of the modules that the developers had in mind. Hence we have continually observed and made videos of teachers engaged in implementing both the draft and finalized versions of the modules. So, now, as we move from the development to the implementation phase of our project we move on with a systematic plan for examining whether the implementation of the modules is going as intended.

Project Background

Our project, the “Engaging Youth Through Engineering Module Study” received funding from the National Science Foundation to develop and test out engineering design modules in middle schools. Initially there were 9 modules that would take place in 6th, 7th, and 8th grades. However, one of the modules that needed extensive revision had to be put on the shelf in order to develop two new modules and revise the others. So, at present, students work on 2 modules in sixth grade, 3 in seventh, and 3 in eighth. The modules cover various topics that link to the Alabama grade level course of study in math and science. Unique to these modules is that they involve work in both math and science classes. The modules share some features of project-based and constructivist based curricula, but also involve some more direct instruction to help support the research and design process. Along with developing the modules, the project involves testing out implementation of the modules across a three year time period to examine the impact of the modules on student learning and attitudes, teacher practices and beliefs, and practices surrounding STEM in the local school district.

Primary to determining whether the modules are working is the notion of fidelity of implementation. Are the modules implemented in ways that the developers view as authentic to the intentions of the developers? Recently, Century, Rudnick, & Freeman (2010) provided a thoughtful discussion of the process of determining fidelity of implementation using a combination of classroom observation, document analysis, and self-report data. They developed a suite of tools for determining fidelity of implementation. These tools, from our perspective were well suited for determining fidelity of implementation of stable fully developed curricula in larger implementations. Our modules are still in the development phase, and although our implementation study has begun, the curriculum is still a work in progress. Further, there are some unique as well as shared elements of the modules, so a more specific system for observing in the classroom is necessary from our perspective. In our case, where we were developing new curricula, understanding of fidelity requires a dialogue between developers, teachers, and evaluators.

In this paper, we focus on the observational component of the fidelity of implementation and discuss the methods we have developed for characterizing the instructional and learning activities
Development of the EYE Classroom Observation Protocol

An important advantage of being part of a project that involves design and development as well as summative evaluation of a program is that opportunity to explore different assessment tools. So, at the beginning of our project, we began looking for tools for doing systematic ratings of classrooms. The first tool we came upon was the Reformed Teaching Observation Protocol or RTOP (Sawada et al., 2002). The RTOP was attractive for a number of reasons: 1. There was good documentation and some online training modules, 2. High scores on the RTOP had been associated with better learning in both mathematics and science, and 3. RTOP could be used to compare teaching in both our project schools and comparison schools.

However, our initial use of the RTOP suggested that it might not capture all we were interested in finding out about implementation of our curriculum. Not all of the lessons focused on reformed teaching methods. Sometimes instruction was not quite as reform-like when students needed to learn procedures or background information they needed to engage in the engineering activity. Further, we realized that the RTOP focused on ratings and we were interested in also describing what was going on in the classroom. After searching out other instruments, we came upon one developed by Lawrenz, Huffman, and Appeldorn (2002). Their instrument kept running tabs on categories of instructional activities engaged in by teachers, cognitive activities engaged in by students, and students’ engagement in five minute increments. Since the instrument was developed as part of an evaluation of mathematics and science curricula, we felt it contained the necessary categories of activities to capture the activities involved in our engineering modules while at the same time being generic enough to use in comparison classrooms.

We made a few adjustments to use with our project. First we changed the intervals to three minutes rather than five. Technically, since we are videotaping, we could do continuous coding, but the three minute interval makes things simpler for our coders. Second, we added a couple of codes to pick up on elements of the design process. For example, while there was an applying procedural knowledge category, we felt that it addressed more routine procedural work and not the application of procedures in the context of designing something. So we added a category there. We added a couple of other categories, but have subsequently dropped them after finding that our coders and others looking at the system could not differentiate them from other existing codes.

After working with coding system, we decided that maybe it would still be valuable to include the RTOP as well. The RTOP has been used a great deal and gives us an anchor for our system that allows us to create an evaluative index of how “reform-like” our instruction appears within and between modules. The two instruments we feel provide us with both a descriptive and evaluative system for observing classrooms. It also provides us with systems that complement each other. For example we believe that coding a lesson using the Lawrenz et al. based coding system helps our coders be more precise in filling out the RTOP.
Returning to fidelity of implementation which is the primary focus of this paper, we felt that the modified Lawrenz et al. instrument would provide with the most information about fidelity of implementation. We should see in our classrooms implementing the modules specific patterns of instructional and student activities. For example, in lessons where students are engaged in design processes, we should see more cooperative learning activities and higher level discussion among students than lecturing. Further, when we code a large number of off task activities, or we see teachers spending too many intervals in logistical activities indicative of a lack of preparation, we can see that the module is not being implemented with fidelity. While it is simple to see the lack of fidelity of implementation when looking at off task behaviors, it is also important to consider the issue of what constitutes the pattern of activities that are indicative of implementation that is consistent with the vision of the developers of the modules. It is to this issue we turn to next.

A Template Matching Approach to Fidelity

During the late 1980s mid 1990s, one of the authors (Van Haneghan) was involved in research exploring a problem solving curriculum called the “Adventures of Jasper Woodbury” developed by the Cognition and Technology Group at Vanderbilt (1994). In a set of case studies, Van Haneghan and Stofflett (1995) found that without assimilating the learning theory underlying the curriculum that teachers often did not implement the problems in ways that were consistent with the vision of developers. With that lesson in mind, we felt that our strategy for examining fidelity of implementation would be to turn to the writers and developers. So we are asking the developers and writers to help develop templates for what implementation should look like using the Lawrenz et al. protocol. It is to an example of this that we turn next.

Method

To illustrate this process we asked the main editor of the series (and one our authors) to take two lessons from “Harnessing the Wind” a module that is in its relatively final form that we have implemented in sixth grade. We gave her the observational coding sheet and she took the module manual and guide to create. We had her develop templates for key lessons in science and mathematics. Given space constraints only the science lesson will be discussed here. Figure 1. Presents the “lesson at a glance”. As can be seen this lesson involves designing and building the wind turbines. Our strategy is to observe key lessons where students are most engaged in the design process (building, testing, collecting and analyzing data). Other lessons will be monitored via self-report forms and reports from project staff involved in the schools with the implementation. This is our second round working through this process. In the first one, the first author in conjunction with project staff developed templates for another module. So we are moving ahead with process and engaging the developers in more depth with this process.
Our templates are basically filled out observation forms that we summarize in two ways to compare with actual classes. First, we examine the % of the time intervals where the various instructional and cognitive activities occurred. Second, we examine the pattern of activities across time in the classroom. These are carried out using graphs generated from the results in Microsoft Excel. We can then visually inspect the patterns of expected activities with those that actually occurred in the lesson in actual classrooms. We could probably do a more sophisticated analysis using a video coding software, but this analysis can be done without the learning curve necessary to master video coding software. However, we are investigating ways to use

Results

Template for the Lesson
Below in Figure 1 are the proportions of time where particular codes are present as envisioned by the series editor. The key for the abbreviated codes is in the Appendix. When discussing codes, I will describe abbreviations as they pop up in the discussion. As can be seen, the lesson has a small proportion of time spent on lecture (L) and Lecture with Discussion (LD). Almost half the segments show students involved in hands on activities (HOA) in cooperative groups (CL) engaged in small group discussions (SGD). There is also a large number of segments with written work (WW) given students are designing and recording information about their design attempts.

Figure 1. Proportions of Time Envisioned by the Series Editor on Various Instructional Activities

Figure 2 presents the cognitive activity codes. As can be seen in Figure 2, the majority of time in the lesson is spent in constructive processes where students are discussing, building, and testing their initial designs. Notice that knowledge representation and both creative and more mundane application of knowledge are required. The most frequent code is the knowledge construction one since the students are working together and testing out ideas. Several of the other codes co-occur with this code. In some ways, if there is knowledge construction, then several of the other codes should co-occur. However, the occurrence of other codes can occur independent of this level of cognitive activity.
Having presented the template developed by the project staff, we now will look at two teachers codes surrounding this lesson in comparison to the developers template. Figures 3 and 4 present the two teachers with their patterns next to the template pattern of the developer/editor. As can be seen, both teachers are not too far away from the pattern indicated by the developer, although both, especially teacher #1 tend to lecture more. Teacher #1 shows lecture in around 60% of the intervals. It can also be seen that Teacher # 1 might be doing some things differently. For example, the template developed shows that

Figure 3. Teacher# 1 in comparison to Template (bars on left are template, bars to right are teacher # 1).

Formative Uses of This Approach

Clarify what developers mean
Help improve the observation system
Data for improving performance of teacher
  Show them the data
Help build better professional development
Help build better curriculum
Helps anchor other measures
  RTOP for Science Lesson 1 – expected to be lower and is lower than Lesson 2
  RTOP for Science Lesson 2- Expected to be higher than Lesson 1 and is higher than Lesson 1.

As Part of Formal Evaluation of Implementation
Can see evidence of strong or weak implementation
Picks up on the necessarily heterogeneous nature of instruction
  Sometimes lecture, or whole class instruction is part of the plan
  RTOP might pick up on the reformed parts of the modules, but not the more intentionally traditional parts

References

Appendix

<table>
<thead>
<tr>
<th>Instructional codes</th>
<th>Cognitive Activity Codes</th>
</tr>
</thead>
<tbody>
<tr>
<td>L= Lecture</td>
<td>Receipt of knowledge = 1</td>
</tr>
<tr>
<td>TIS= Teacher/faculty interacting w/ student(s)</td>
<td>Application of procedural knowledge = 2</td>
</tr>
<tr>
<td>PM= Problem Modeling</td>
<td>Application of knowledge, creative = 3</td>
</tr>
<tr>
<td>HOA= Hands-on activity</td>
<td>Knowledge representation = 4</td>
</tr>
<tr>
<td>SP = Student presentation</td>
<td>Knowledge construction/Constructive Discourse = 5</td>
</tr>
<tr>
<td>SGD = Small group discussion, 2 or more</td>
<td>Other (Please describe) = 0</td>
</tr>
<tr>
<td>LWD= Lecture w/ discussion</td>
<td></td>
</tr>
<tr>
<td>I= Interruption</td>
<td></td>
</tr>
<tr>
<td>D = Demonstration w/ equipment</td>
<td></td>
</tr>
<tr>
<td>OTH=Other (describe)</td>
<td></td>
</tr>
<tr>
<td>UT = Digital media/technology</td>
<td></td>
</tr>
<tr>
<td>WW=Writing work, add SGD if in groups</td>
<td></td>
</tr>
<tr>
<td>CD =Class discussion, all</td>
<td></td>
</tr>
<tr>
<td>A = Assessment (describe)</td>
<td></td>
</tr>
<tr>
<td>AD</td>
<td>Administrative tasks</td>
</tr>
<tr>
<td>----</td>
<td>----------------------</td>
</tr>
<tr>
<td>RSW</td>
<td>Reading seat work, add SGD if in groups</td>
</tr>
<tr>
<td>OOC</td>
<td>Out of class experience</td>
</tr>
<tr>
<td>CL</td>
<td>Cooperative learning (roles)</td>
</tr>
<tr>
<td>LC</td>
<td>Learning center/station</td>
</tr>
</tbody>
</table>