Bringing STEM Graduate Fellows into Rural Middle Schools: 
First-Year Results of the NSF-Funded 
Indiana Interdisciplinary GK-12 Project using a Mixed-Methods 
Approach to Program Evaluation

Abstract

The overall goal of this paper is to share results from the first year of the NSF-funded, Interdisciplinary GK-12 project: Bringing Authentic Problem Solving in STEM to Rural Middle Schools while providing an example of our mixed-methods approach to program evaluation. The GK-12 project pairs STEM doctoral fellows with middle school science teachers to achieve four project goals: 1) Enhance the quality of STEM education in middle schools; 2) Provide training and mentoring for graduate student fellows; 3) Provide professional development for teachers; and 4) Foster mutually beneficial partnerships between the university and area school corporations. Participants are 8 doctoral fellows, 8 middle-school teachers, and 414 middle-school students. Data have been collected on several measures of teaching self-efficacy, science and teaching concept formation, and perceptions of scientists to determine if differences exist before and after program participation for graduate fellows, teachers, and middle school students. Both quantitative and qualitative data have been collected for each program goal to triangulate findings and enhance interpretation of the results. Preliminary findings suggest that the GK-12 program is achieving the desired outcomes for fellows, teachers, and students. Results also indicate the benefits of using a mixed-methods approach for program evaluation.
In his opening address to the Center for the Integration of Research, Teaching, and Learning (November 5, 2003), then Director of the National Science Foundation, Dr. Joseph Bordogna, offered the following observations about the importance of integrating science education across all levels of learning.

“Graduate education simply cannot be considered in isolation. As we weigh the balance between continuity and reform, let us consider how we can work towards a system of education that will be a seamless route of advancement for students from K-12 through post-doctoral levels. There is one particular divide that I would like to highlight—the divide between K-12 “teachers” on the one hand, and college and university "faculty" on the other. We need to move toward recognizing the entire spectrum of our educational professionals as "faculty." Early learning is as important to individual development, and to long-term social progress, as is a higher degree. We will only be able to design seamless learning paths for students when we recognize our common ground as faculty in a seamless educational community.”

More recently, NSF Director, Arden Bement, presented the concept of the “civic scientist” in his Inaugural Civic Scientist Distinguished Lecture at Rice University (November 9, 2006). Dr. Bement described the responsibility of scientists to help “advance the knowledge base of the nation’s science and math teachers and their students” through active engagement in classrooms and the design of hands-on learning experiences for K-12 students.

As we face a growing crisis of preparedness for the scientific challenges of the 21st century, leaders such as Drs. Bordogna and Bement have begun to conceptualize and fund programs that redefine the landscape of STEM education. “Integration,” “communication,” “civic responsibility,” “experiential learning,” “collaboration” are words that describe the new terrain of science, math, technology, and engineering education.
In 1999, the National Science Foundation (NSF) began funding the GK-12 Education Program to support teaching fellowships for graduate students in STEM disciplines (science, technology, engineering, and math). Based upon a comprehensive NSF model of graduate education (National Science Foundation, 2006) the program was designed to integrate teaching and learning through long-term, systematic interactions between K-12 teachers and graduate research scientists in the context of the K-12 classroom. It was reasoned that the two-way interaction between teachers and graduate scientists would result in benefits to fellows who would gain experience in communicating science concepts, teachers who will infuse new science concepts into their curriculum, and students who would receive additional mentoring from a cadre of research scientists in the classroom. The originators of the GK-12 initiative hoped to provide graduate students with skills that would better prepare them for the challenges of scientific careers in the 21st century while providing deeper understanding of the nuances of their scientific research. “The ultimate goal of the NSF GK-12 program is to improve the awareness and communication skills of the next generation of scientists and engineers by providing them with direct teaching experience in grades K to 12, and in the process, improve the quality of science education at the primary and secondary level in the United States” (NSF, 2006).

To date NSF has funded 217 GK-12 initiatives with projects in nearly every state (NSF, 2006). These projects have applied the basic GK-12 philosophy to a variety of grade levels, a variety of science, mathematics, technology, and engineering disciplines, and with a variety of support structures. The common element of all programs, however, is the inclusion and support of graduate student fellows in intensive school-based environments in which participants are able to take meaningful roles in content and pedagogical experiences.
The Indiana Interdisciplinary GK-12 Program is one example of the efforts of NSF to extend this model to rural middle school settings. Originating in the Purdue University Discovery Learning Center, the Indiana GK-12 has matched eight teachers from four middle schools with eight graduate fellows from a variety of science and mathematics disciplines. Two fellows and two teachers from each school interact as a team to discuss pedagogical and curricular issues. Additionally, fellows meet biweekly with the GK-12 support team at Purdue to address progress, insights, and challenges of their experience. The GK-12 is organized around an interdisciplinary team model designed to approximate current trends in scientific research while capitalizing on authentic problem solving as a mechanism for developing interest and motivation for STEM study in K-12 students.

This purpose of this paper is to share one GK-12 evaluation model, discuss our assessment instruments and design, and communicate initial findings from our first year of implementation.

**Developing an Evaluation Model**

**The Importance of Articulating Program Theory**

Too often program development and evaluation focuses on the relationship between interventions and outcomes. Chen (2005) refers to this common approach as an input-output or “black box” evaluation: things go in and things come out, but what happens in between is a mystery. While useful in identifying program merits, the black box evaluation cannot capture the “transformation processes that turn interventions into outcomes” (p. 231).

Theory-based evaluation provides both an understanding of how transformational processes can be used to understand and improve a program and a description of the attributes of the program. The added value of theory-based evaluation is an understanding of how and why change occurs as
well as knowledge of what and how much change has taken place. Chen (2005) describes three advantages of evaluation based on theory: 1) by understanding underlying transformational connections, theory driven evaluations provide information for program improvement while contributing to the broader science, in general, 2) theory-driven evaluation has greater construct validity because it examines causal mechanisms that underlie the program, and 3) theory-driven evaluation can increase internal validity by requiring the specification of program theory.

When we engage in the development of a program such as the Indiana Interdisciplinary GK-12, there are underlying assumptions about the relationships between program means (resources, activities) and ends (outcomes, impacts). The articulation of these relationships constitutes what is referred to as program theory. “The construction of a means-ends hierarchy for a program constitutes a comprehensive description of the program’s model” (Patton, 2000, p. 217). Logic models provide one approach to representing program theory.

**Logic Models**

All logic models provide a graphical representation of program elements and the logical, temporal relationship between the components. While the model can be represented in a variety of ways, the logic model framework proposed by the Kellogg Foundation (2004) is both easily understood and widely used. The logic model consists of the following elements:

- **Inputs/Resources** – Inputs are program resources and typically include such things as funding, personnel, facilities, materials, time, consultants, transportation and other resources that are specific to a given program. Program implementation is dependent on adequate inputs.

- **Activities/Processes** - Activities are the action components of the program and include transformational processes aimed at attaining desired results for the participants. Examples of activities include training sessions, workshops, seminars, and team building exercises.

- **Outputs** – Outputs include tangible evidence of program delivery such as number/type of participants, number of hours completed, number of staff recruited and trained, and products or materials developed (e.g. handbook, training materials, curricula). Outputs are tied to and
dependent on the program activities and are circumscribed by the inputs supporting the program.

• Outcomes – Outcomes are directly related to the outcome objectives of the program and describe the anticipated results of program implementation. Outcomes can be considered in terms of proximity of impact with short-term objectives relating primarily to learning (such as knowledge, awareness, attitudes, skills) and medium-term objectives relate to anticipated changes in behavior (e.g. increases in positive behaviors and practices, decreases in nonproductive behaviors).

• Impacts - Impacts are those outcomes that might be expected long term and include such things as institutional changes, social changes, and more generalized change that might not be immediately apparent after the implementation of the program.

• Context/Environment – Important in the representation of a program is consideration of the factors that may interfere with the planned delivery of the program. Anticipating these factors may allow for initial program modification. If it is not possible to work around contextual constraints, the understanding of how these factors influence program outcomes can be invaluable in extending program theory and increasing program benefits through formative evaluation processes.

• Assumptions – Assumptions provide the basis for inferring that program elements will result in anticipated outcomes or impacts. Assumptions can be derived from developmental, social, or learning theory, empirical evidence or espoused theories.

As the GK-12 project was being conceptualized, it was important to clearly outline the project goals and objectives and determine how identified processes and outcomes would be measured. A logic model was created to represent the relationship between broad goals, project resources, planned activities, and desired outcomes and impacts. Based on recommendations from NSF and the Kellogg Foundation (2004), the GK-12 project evaluators worked closely with the project developers to ensure model coherence and measurable program outcomes.

**Project Goals**

As noted above, the overall purpose of the GK-12 initiative is to improve STEM education by providing 1) professional development, curriculum development opportunities, and content matter support to teachers, 2) interdisciplinary, inquiry-based curriculum and science role models for students, and 3) enhanced communication, teaching abilities, and understanding
of the role of researchers in K-12 education for STEM graduate students and their faculty mentors (NSF, 2006). The four project goals for the Indiana Interdisciplinary GK-12 include:

1) To enhance the quality of STEM education in middle schools, particularly those within small town or rural communities, and with a multicultural and/or lower socioeconomic component, thus increasing student STEM knowledge and interest;

2) To provide training and mentoring for graduate student fellows, including experience in middle school classrooms, and enhancements to develop their ability as future STEM faculty or industry professionals;

3) To provide professional development for middle school teachers, enhancing offerings and inquiry-based education in STEM disciplines; and

4) To foster ongoing, mutually beneficial partnerships between the university and area school corporations to continually enhance STEM education.

The following figure represents the logic model for goals 1 (K-12 students), 2 (graduate fellows), and 3 (teachers). Because goal 4 (partnerships) is not discussed in this paper, the logic model does not include information about this goal.
Figure 1: Logic Model - Indiana Interdisciplinary GK-12

Goal 1: Enhance the quality of STEM education in middle schools, thus increasing student STEM knowledge and interest.

Goal 2: Provide training and mentoring for graduate student fellows to develop their ability as future STEM faculty or industry professionals.

Goal 3: Provide professional development for middle school teachers, enhancing offerings and inquiry-based education in STEM disciplines.

<table>
<thead>
<tr>
<th>Inputs/Resources →</th>
<th>Activities Processes →</th>
<th>Outputs→</th>
<th>Outcomes→</th>
<th>Outcome Measures→</th>
<th>Impacts</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSF Funding</td>
<td>Maymester Course</td>
<td># fellows</td>
<td>Increased teaching self-efficacy</td>
<td>STEBI-B Attitudes and Beliefs Survey (5/06 &amp; 5/07)</td>
<td>Cadre of scientists who are able to communicate scientific understandings to a broad audience (Fellows)</td>
</tr>
<tr>
<td>Purdue Faculty</td>
<td>Summer Workshops</td>
<td># teachers</td>
<td>Broadened perceptions of “scientist”</td>
<td>Draw-A-Scientist (5/06 &amp; 5/07)</td>
<td>Teachers who are confident and effective in engaging learners with current scientific content (Teachers)</td>
</tr>
<tr>
<td>Professional Staff</td>
<td>Orientation to Middle School</td>
<td># learning teams</td>
<td>Increased knowledge of pedagogy</td>
<td>Concept maps Program Survey (5/06 &amp; 5/07)</td>
<td>Increased K-12 STEM learning and motivation (K-12 Students)</td>
</tr>
<tr>
<td>Graduate Research Assistants</td>
<td>Classroom Participation</td>
<td># schools</td>
<td>Increased knowledge of scientific inquiry</td>
<td>Observations (twice/semester)</td>
<td></td>
</tr>
<tr>
<td>Purdue Facilities</td>
<td>Biweekly Fellows Meetings</td>
<td># hours engaged</td>
<td>Increased awareness of middle school students and environments</td>
<td>Concept maps Program Survey (5/06 &amp; 5/07)</td>
<td></td>
</tr>
<tr>
<td>Middle School Facilities</td>
<td>Teacher-Fellow-GK-12 Team Meetings</td>
<td># sessions</td>
<td>Increased interest in STEM careers</td>
<td>Journal entries (ongoing)</td>
<td></td>
</tr>
<tr>
<td>Technology &amp; Curricular Materials</td>
<td>Graduate Fellow Mentoring</td>
<td># meetings</td>
<td></td>
<td>Interviews Open-ended questions (5/06 &amp; 5/07)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td># K-12 students impacted</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td># lesson plans</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Model Assumptions
Long-term partnerships between teachers and graduate scientists enhance the quality and success of STEM education on all levels.

Contextual Factors
Tension between research and teaching demands of graduate fellows. Personality and philosophical “fit” between teacher and fellow. School factors (distance, scheduling, resources, etc).
Goal 1: Enhance the quality of STEM education in middle schools, thus increasing student STEM knowledge and interest

Goal 1 of the Indiana Interdisciplinary GK-12 is focused on outcomes involving *middle school students*. These students had a visiting GK-12 scientist in their math or science classrooms for the year for approximately 10 hours per week. Constructs of interest include increased motivation to pursue science careers or fields of study, confidence in math and science abilities, broadened perceptions of scientists, and greater STEM knowledge. These constructs were measured by quantitative surveys, open-ended questions, and classroom observations, which will be described in detail in the Methodology section.

Goal 2: Provide training and mentoring for graduate student fellows

Goal 2 is linked to activities and outcomes involving the *graduate student fellows*. These participants have been selected on the basis of demonstrated excellence in their fields of STEM study and receive an annual stipend. They had the opportunity to engage in learning teams and have participated in an orientation to middle school environments and workshops/classes designed to develop an understanding of STEM curriculum development and pedagogy. The fellows also participated in the middle school classrooms at least 10 hours per week during which they gain experience with both teaching and the development of instructional materials. Variables of interest include fellows’ knowledge of pedagogy, understanding of complex learning environments, increased ability to communicate scientific concepts, better understanding of scientific inquiry, and personal confidence in the classroom. As noted in the logic model, these variables are being evaluated through semi-structured interviews, questionnaires, surveys, drawings, and concept maps. These instruments and procedures are described in detail in the following section.
Goal 3: Provide professional development for middle school teachers

Goal 3 is focused on the activities that relate to participating middle school teachers. There is considerable overlap between the outcomes for goals 1 and 2 since fellows and teachers participate together in learning teams, summer workshops, development of educational modules, and classroom instructions. It is expected that participation in the project will influence the STEM content knowledge of middle school teachers and the delivery of that knowledge in classroom settings. The project workshops and class visits should also have a positive impact on teacher attitudes toward teaching, in particular, inquiry-based learning methods. These variables will be measured through classroom observations, teacher interviews, teacher surveys, concept maps, drawings, and open-ended questionnaires. Again, these instruments and assessment procedures are described in detail in the Methodology section below.

Methodology

Participants

Participants consisted of three groups: eight graduate fellows, eight middle school teachers, and 414 seventh- and eighth-grade middle school students.

Fellows. There were four female and four male fellows. The majority of the fellows were Caucasian. All of the graduate fellows were doctoral students in a math or science-related field. Fellows’ previous teaching experience, including teaching assistantships, ranged from two to seven years. Most responded that they planned to find a primarily teaching position at a university or college.

Teachers. There was one male teacher and seven female teachers, all of whom were Caucasian. Teachers’ educational backgrounds ranged from a Bachelors degree to a Ph.D. degree
with most responding that they planned to earn additional degrees. Teaching experience ranged from two to 31 years.

Students. Students were from four local middle schools, which varied with regards to ethnic makeup and socioeconomic level. Fifty-one percent of students were female, 48% were male, and 1% were not reported. The majority of students were Caucasian (76%), 8% were Hispanic, 3% were African American, 2% were Asian American, and 11% were not reported.

Design

This evaluation uses a mixed-methods design, incorporating both qualitative and quantitative measures for each participant group for each outcome.

Instruments and Procedures

Science Teaching Efficacy Belief Instrument-Pre-service (STEBI-B)

The Science Teaching Efficacy Belief Instrument-Pre-service (STEBI-B; Enochs & Riggs, 1990), which is based upon the social learning theory of Albert Bandura (1977), was developed to measure the self-efficacy of preservice teachers as a predictor of science teaching behavior. The STEBI-B consists of 23 items that comprise two subscales: Personal Science Teaching Efficacy (PSTE) and Science Teaching Outcome Expectancy (STOE). Items use a 5-point Likert scale. Internal consistency reliability for the PSTE subscale is high (α = 0.90) and reliability for the STOE subscale is moderate with α = 0.76 (Enochs & Riggs, 1990). Graduate fellows responded to the STEBI-B before and after the GK-12 teaching experience.

Concept Maps

A concept map is a visual representation of the relationship between concepts (Novak & Gowin, 1984; Ruiz-Primo, 2004). Concept maps were developed by Novak and Gowin (1984) as a tool for evaluating understanding of conceptual knowledge and to promote meaningful
learning. GK-12 fellows and teachers were asked to create two concept maps using a list of words (concepts) related to the teaching and learning of science. As a post-measure, the maps were returned and fellows and teachers were instructed to make revisions as necessary.

**Draw-a-Scientist**

The Draw-A-Scientist Checklist (DAST-C) was developed by Chambers (1983) to assess students’ perceptions of scientists by using a checklist of stereotypical attributes to score drawings of scientists. The DAST-C consists of 16 stereotypes: the scientist wearing a lab coat, eyeglasses, having facial hair, displaying research symbols, displaying knowledge symbols, technology representation, relevant captions, male gender only, Caucasian only, middle aged/elderly, mythic stereotypes, indications of secrecy, working in a lab, indications of danger, and clothing/expressions. Fellows, teachers, and students were asked to draw a scientist at the beginning and end of the GK-12 program.

**Pre and Post Program Survey**

The Pre/Post Program Survey was developed for the purpose of assessing teachers’ and fellows’ confidence in activities related to teaching and their beliefs about the importance of these activities to teaching. Versions of the survey differed slightly for teachers and fellows. The fellows’ Program Survey consisted of 52 items on a Likert-type scale and concentrated on issues such as teaching confidence and the importance of communicating scientific concepts to a broad audience. The teachers’ Program Survey contained 49 items on a Likert-type scale and focused on STEM teaching self-efficacy and the importance of teaching STEM concepts.

**Qualitative Measures**

Qualitative data were collected throughout the year for fellows and teachers. An example is a series of reflective questions that fellows completed biweekly. The purpose of these
questions is to prompt the fellow to reflect more deeply about his or her teaching experiences.

These questions are different each time; for example, one set of questions asked:

1) Describe similarities/differences between your middle school experience and the GK-12 middle school.

2) What are the main challenges facing middle school teachers, and how can they be resolved?

3) Have you noticed differences in learning styles based on differences in students? What are they, and does the teacher address them? If so, how? If not, why?

Another source of qualitative data are open-ended questions that are part of the Pre/Post Program Survey, which ask teachers and fellows about what they gain/ed from the project and what challenges they face/d. Another question asks the fellows about what they contribute to students’ learning. There were also several open-ended questions for middle school students, which focused on students’ career interests and enjoyment of math and science.

Classroom observations of the teachers’ and fellows’ interactions with one another and with the students took place throughout the school year. A classroom observation form was used to help focus the observations and extensive notes were recorded in observation journals. Observations were made by two observers, one of whom was frequently in the classrooms and served as the program liaison. The second observer served as an evaluation assistant and visited each school one or two times mid-year for the purpose of comparing observations and gaining perspective into the classroom dynamics.

*Inventory of Children’s Activities – Revised (ICA-R)*

The Inventory of Children’s Activities – Revised (ICA-R; Tracey & Ward, 1998) was designed to measure children’s interests and self-efficacy. The ICA-R consists of two sections of 30 activities in which children commonly engage. The first section asks students to rate how
confident they are at performing the activities (1=don’t like at all, 5=like a lot) followed by asking students how competent they believe themselves to be for each of the activities (1=very bad at, 5=very good at). The ICA-R has shown good internal consistency and test-retest reliability (Tracey & Ward, 1998). The ICA-R was administered to the middle school students before and after the program to determine changes in career interests and levels of perceived competence in these interests.

**Attitude towards Math and Science Scales**

Students were asked to complete a modified version of the Attitude towards Math and Science Scales (Erkut & Marx, 2005) at the beginning and end of the school year. The survey consists of 23 items for which students are asked to indicate their level of agreement with positive or negative statements about math and science. The instrument was piloted in several eighth grade classrooms and shows good internal consistency reliability, with Cronbach’s $\alpha = 0.87$ and 0.86 for the pre and post Attitudes towards Math Scale, respectively, and $\alpha = 0.87$ and $\alpha = 0.88$ for the pre and post Attitudes towards Science Scale, respectively (Erkut & Marx, 2005).

**Confidence in Science and Math Abilities Survey**

The Confidence in Science and Math Abilities Survey was developed to assess middle school students’ confidence in their ability to perform tasks that are based on the 7th and 8th grade science and mathematics standards. Students are asked to indicate their confidence in a total of 28 items on a five-point scale (1=Not At All Confident, 5=Very Confident). There are two versions; one for 7th grade standards and one for 8th grade standards, which contain the same number of items for each. The survey was administered as a pre – post measure.
Results

Anticipated Student Outcomes

There were several anticipated outcomes for middle school students based on the broad goal of increasing the quality of STEM education with regards to STEM knowledge and interest. These outcomes include motivation to pursue STEM careers or studies, broadened perceptions of scientists, interdisciplinary learning opportunities, increased confidence in performing math and science-related activities, and increased factual knowledge about STEM subjects. The following is an overview of the results of selected student outcomes.

Broadened Perceptions of Scientists

The Draw-a-Scientist Checklist (DAST-C) was scored according to the number of stereotypes that were drawn. Positive traits, which include a smile and regular clothing, were also scored.

Using a Wilcoxon signed-rank test, results show that overall, students drew significantly less stereotypical images of scientists at the completion of the GK-12 program than the beginning ($p < .05$). Additionally, students drew significantly more positive traits ($p < .05$) at the end of the year than at the beginning.

In addition, an open-ended question that asked about students’ future career interests was analyzed to determine if the DAST-C findings are corroborated. It was expected that students would not only list a greater number of STEM-related careers at the end of the GK-12 program but also list a wider variety of STEM careers as their perceptions of what scientists do are broadened.

Because of the number of responses and non-matching student identification numbers, a random sample of 50% of students on the pre-survey and 50% of students on the post-survey was
chosen for analysis. The open-ended question was scored according to the percentage of students choosing a career in science, technology, engineering, math, a non-STEM field, or unsure of a career path. Gains were found from pre to post responses for students’ career interests in science, engineering, and math. There were also anticipated decreases in the percentage of students who listed a non-STEM field or an unsure response from the pre- to the post-survey. An unexpected outcome was a decrease in the percentage of students who listed a career in technology from pre to post survey. Table 1 displays the number and percentage of student responses for each career interest area from the pre- to post-survey.

Table 1

<table>
<thead>
<tr>
<th>Area of Career Interests</th>
<th>Percentage of Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre</td>
</tr>
<tr>
<td></td>
<td>n=273</td>
</tr>
<tr>
<td>Science</td>
<td>29%</td>
</tr>
<tr>
<td>Technology</td>
<td>8%</td>
</tr>
<tr>
<td>Engineering</td>
<td>4%</td>
</tr>
<tr>
<td>Mathematics</td>
<td>0%</td>
</tr>
<tr>
<td>Non-STEM Area</td>
<td>49%</td>
</tr>
<tr>
<td>Don’t Know/Unsure</td>
<td>10%</td>
</tr>
<tr>
<td>Total STEM Interests</td>
<td>41%</td>
</tr>
</tbody>
</table>

The analysis of the DAST-C revealed that students were able to envision a less stereotypical version of a scientist. This is shown by the significant decrease in stereotypical
drawings as well as a significant increase in positive attributes. The qualitative analysis of the open-ended survey question supports this finding because the majority of students listed a greater percentage of STEM-related careers on the post-survey. It is possible that by picturing scientists as less stereotypical that students are better able to imagine themselves pursuing a STEM-related career, which is a goal of the GK-12 program.

**Interdisciplinary Learning Opportunities**

One outcome of the GK-12 program is to increase interdisciplinary learning in math and science classrooms. For example, a mathematics doctoral fellow may use science examples in a math classroom to explain math concepts. Evidence for this outcome includes teacher and fellow focus groups, Program Survey open-ended questions, and fellows’ reflective questions.

Teachers reported positive outcomes in regards to the GK-12 program creating an interdisciplinary learning environment. One teacher explains:

“[I have gained] confidence in … my ability to ‘do science’ with the kids in a ‘math’ environment, my ability to incorporate new and innovative technologies into the classroom – both for my teaching and for their activities in learning, my ability to create problem solving students (or at least more so than I think I was in the past), confidence to approach learning in different ways, etc. Maybe the most important thing I’ve gained is a sense that it is so important to connect all of what we do in math to the real world – specifically the scientific community – so that the students have a broader selection of futures from this point forward. I think we’ve changed the thought processes of several of the students and many of them are interested in the stuff that we did with them and the things that we introduced to that they may, indeed, pursue these things at a higher level as they get older.”

Other teachers say:

“I also was more aware of the science curriculum and how to tie it in to my math class.”

“It was really eye-opening and has changed the way that I will teach. Incorporating science aspects into my classroom seems so much easier and relevant now. I wondered how this was all going to work in the beginning – me
being math, the program seemingly aimed towards science…but it really is all one in the same when we get down to it. I’m not sure I really realized it before.”

Some teachers reported that this interdisciplinary focus was a challenge, explaining:

“Having a fellow that is more math oriented and not science [was challenging], so sometimes connecting the science was a problem.”

“I do think this program heavily favors science over math even though it was pitched as both math and science.”

**Anticipated Fellow Outcomes**

Anticipated outcomes for the graduate fellows include increased teaching self-efficacy, increased knowledge of pedagogy, and increased ability to communicate science concepts and research.

**Teaching Self-Efficacy and Knowledge of Pedagogy**

The pre and post STEBI-B responses were analyzed to determine if there were significant differences from before and after the GK-12 project. Using a Wilcoxon signed-ranks test, there were no significant differences between the pre and post STEBI-B (\( p = .24 \)) and (\( p = .53 \)), respectively, for the two subscales. The fellows’ mean ranks on the pre STEBI-B are: (Subscale 1 = 3.20) and (Subscale 2 = 3.38) and on the post STEBI-B the mean ranks are: (Subscale 1 = 5.00) and (Subscale 2 = 3.75).

Although STEBI-B scores improved over the course of the year, statistical analyses of pre and post program STEBI-B scores were not significant. Post program concept maps helped us understand changes in inquiry frameworks, suggesting a richer and more nuanced development than revealed in the STEBI-B, alone.

Research has shown that concept maps may capture the effects of instruction even where standardized tests do not (Wallace & Mintzes, 1990). Therefore, GK-12 fellows were asked to create a concept map using a list of words (concepts) related to the teaching and learning of
science. As a post-measure, the maps were returned and participants were instructed to make revisions as necessary. Novak’s scoring criteria (Novak and Gowin, 1984) was used to score concept maps according to the number of levels in the hierarchy of concepts, the number of links, the validity of the links and concepts, and the cross-links between different portions of the map. Concept maps that are mostly hierarchical with no cross-links suggest a lack of awareness of the interconnections between the concepts (Turns, Atman, & Adams, 2000). For this reason, cross-links are scored the highest number of points (10). The number of levels are scored five points each, the number of links are given one point each, and additional examples are one point each.

As shown in Table 2, there were gains in fellows’ concept map scores from pre to post. Although cross-links were not used in several fellows’ concept maps, there was also an overall gain in the number of cross-links that were used from pre to post concept map.

Table 2

Concept Map 1: Teaching and Learning

<table>
<thead>
<tr>
<th>Fellow</th>
<th>Total Score</th>
<th>Gains</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre</td>
<td>Post</td>
</tr>
<tr>
<td>1</td>
<td>79</td>
<td>106</td>
</tr>
<tr>
<td>2</td>
<td>54</td>
<td>114</td>
</tr>
<tr>
<td>3</td>
<td>117</td>
<td>117</td>
</tr>
<tr>
<td>4</td>
<td>31</td>
<td>74</td>
</tr>
<tr>
<td>5</td>
<td>68</td>
<td>79</td>
</tr>
<tr>
<td>6</td>
<td>50</td>
<td>59</td>
</tr>
<tr>
<td>Total</td>
<td>399</td>
<td>549</td>
</tr>
</tbody>
</table>
Additional evidence for increased teaching self-efficacy and knowledge of pedagogy was revealed in the fellows’ focus group. Several comments showed how fellows became aware of the differences in teaching styles at the middle school level versus the college level and how they believe this will impact their future teaching at the college level. For example, two of the fellows said:

“And for me it’s been really hard to come out of the college mindset, where in college it’s not your problem it’s their problem. Here it is your problem if they cannot understand what you’re trying to give them to do or trying to get them to do. And so it’s just interesting for me to go back and forth between the much more personal, compassionate mindset to the college mindset when you go back to campus. There’s just a lot more compassion and a lot more flexibility, and there should be, in the middle school classroom than what we normally do and give people here at Purdue or in college.”

“I just think that what we learn about individuality and giving special attention in the middle school could definitely be applied to college students and could help a lot of college students without hand holding. I don’t know, I think there’s a lot to be learned for compassion in teaching.”

Classroom observation notes also reveal that teachers are satisfied with the fellows’ integration into the classroom. One of these field notes states:

Observation notes: [The teacher] is happy with the fellow’s performance – the fellow has integrated into the classroom really well, helping with curriculum design (e.g., enhancing the questions on quizzes) and the fellow has also taken the initiative to request the opportunity to “lead” classes in the afternoon, after having had the opportunity to observe [the teacher] in the morning.

Finally, the fellows’ understanding of inquiry-based learning is revealed in their responses to an open-ended item, stating: Discuss the importance of inquiry methods in teaching about the nature of science.

The fellows provided a variety of responses to this question with the majority of fellows responding that students are able to construct their own understanding of the topic, which
promotes critical thinking skills. Table 3 shows the number of fellows who wrote about each of the listed topics.

Table 3

*Main Topics and Number of Fellows who Wrote about Each Topic*

<table>
<thead>
<tr>
<th>Main Topics</th>
<th>No. Fellows Listing Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Students construct their own understanding/test their beliefs/critical thinking</td>
<td>4</td>
</tr>
<tr>
<td>Lasting effect of learning</td>
<td>1</td>
</tr>
<tr>
<td>Enables correct scientific ideas/eliminates misconceptions</td>
<td>2</td>
</tr>
<tr>
<td>Forces students to recognize their prior knowledge</td>
<td>1</td>
</tr>
<tr>
<td>Engages students</td>
<td>1</td>
</tr>
<tr>
<td>Teacher and students more aware of what they think</td>
<td>2</td>
</tr>
<tr>
<td>Generates and sparks curiosity</td>
<td>2</td>
</tr>
<tr>
<td>Students actively seek explanations to real-world problems/ experimentation</td>
<td>2</td>
</tr>
<tr>
<td>Achieve a greater understanding</td>
<td>1</td>
</tr>
<tr>
<td>Students see how the problem is important to everyday life</td>
<td>1</td>
</tr>
<tr>
<td>Social interaction/teamwork</td>
<td>1</td>
</tr>
<tr>
<td>Use interdisciplinary fields</td>
<td>1</td>
</tr>
</tbody>
</table>

**Communication of Science Concepts**

One of the main learning outcomes for the fellows was an increased ability to communicate to others about their field of study. Vignettes taken from the fellows’ focus group reveal these changes:

“I think it has helped me just with communication in general. Communication with a wide variety of people, starting with teachers, because interpersonal skills have become very important with this and understanding how to communicate with your teacher in a way that you are not sounding burdensome or overwhelming to them. But especially communication with the students, just how do you communicate a very complicated concept to a 12 year old.”

“I’ve noticed it more in when I present my research or present to other colleagues. So, my presentation skills outside of GK12 have improved. Am I loud in presentations I give, in talks I give, I just, in integrating more of the I guess you’d
call it GK12 teaching techniques and those other lessons, and I’ve seen personal growth in that way.”

“Clearly this year, this semester, I’ve gone on a bunch of job interviews so I’ve had to take my research and present it to a different audience. So even though I haven’t really talked about my research to the students or the teachers, I still have to communicate science concepts to an audience who doesn’t know anything about it. In a lot of these interviews I wasn’t talking just to the search committee, I was talking to faculty from all over the college or university, not just in earth science. And also just talking to intro level students so I think it has really helped me as a presenter. And on the most recent interviews I went on I kind of had the mindset of I was talking to these 8th graders, and those were my more successful interviews.”

“I think it just brought me out of the I’m the scientist/expert, I’m going to talk in my language, I’m going to try and blow you away, you know, and that doesn’t work.”

In addition to the focus group, evidence for increased communication of science concepts was found in the fellows’ response to one Program Survey question. This question asked: I am confident in my ability to explain scientific ideas to non-expert audiences (p = .317). While this item was not significant, there was an increase in overall scores from pre to post survey.

Although STEBI-B scores and the Post Program item score improved over the course of the year, statistical analyses of pre and post gains were not significant. It was speculated that more change occurred for the fellows than was revealed in these comparisons, which was revealed in fellows’ concept maps.

**Anticipated Teacher Outcomes**

The main outcomes that were anticipated for the teachers was increased knowledge of pedagogy (e.g., inquiry-based learning), increased knowledge of scientific inquiry, increased knowledge of subject matter, and broadened perceptions of scientists.
Knowledge of Pedagogy

Focus group vignettes from the teachers reveal that teachers broadened their teaching styles and learned new pedagogical techniques. For example:

“I appreciated their [fellows’] approach and I learned to be, I don’t think I was textbook dependent, but I definitely used that as my main resource and the fellows didn’t even acknowledge the textbook. So I learned kind of a way to think of a big idea, and think of how you want to teach it, and how you want to show that, and demonstrate it and bring ideas in, and cover standards and not really worry about if it’s in the book or not or how the book explains it. I think science, we barely touched our textbook all year.”

“I taught for 24 years and you kind of get in your own routine and it’s like there’s the book, you know, let’s do it. And I need to step away from that more and this has helped that a little bit. It was a start. I need to do more.”

“Well I think the neat thing about this is it gives us another…I mean we’re all…it’s all ISTEP driven, I mean we’re still focused on wanting these kids to do well but it’s nice to be able to see some other ways of getting it done.”

Knowledge of Scientific Inquiry

During the summer before the classroom experience began, the fellows were required to complete a 1-credit course about educational theory. Inquiry-based learning was one of the main topics and fellows were strongly encouraged to implement inquiry techniques into their middle school classroom lessons. It was expected that teachers would subsequently gain knowledge of inquiry learning by observing the fellows in the classroom. It appeared that teachers grasped the benefits of inquiry, saying:

“I do many hands-on, inquiry type investigations and brain-storming with [the fellow] on new and different ways to approach a topic...”

“...Something that [fellow] and I talked about a lot is that by doing the projects and having the kids do some more inquiry based things, and just projects where they are really thinking, their overall problem-solving skills and perseverance has increased a lot in my classroom.”

“This one day we were discussing a problem in algebra...and I thought, okay, let’s just ask leading questions. And it was a wonderful lesson and nobody forgot.
It was all student generated. I didn’t give answers, I didn’t just show them how to do it, I let them make mistakes… And I found that I am doing that more. I’ve gone through all kinds of enhancements for classrooms; three-second wait and positive reinforcement, but for some reason I just didn’t do more questions you know like two or three. So to me, I think I got something out of this.”

“I was able…something that [fellow] and I talked about a lot is that by doing the projects and having the kids do some more inquiry based things…and just projects where they are really thinking…their overall problem solving skills and perseverance has increased a lot in my classroom. I feel like their problem solving skills overall have increased. Those are the standards that I’ve always had a really hard time with, how do you really hit these problem-solving standards. And I feel like with every project we do we look up at those standards and we’re like yeah we’re definitely working on these. So I don’t feel like I have to find specific activities that hit problem solving. I feel like they’re just automatically learning them.”

**Knowledge of Subject Matter**

Some teachers said that the GK-12 experience allowed them to gain a deeper understanding of their subject area. For example, several teachers said:

“I knew coming into this that biology was my weak area and I feel like I have more depth now. It’s not that I necessarily learned anything new, I just understand it better.”

“I felt like that there were some physics concepts that…there was kind of a number of things…like [fellow], he knows that…he was very philosophical and I tend to be too, so we could argue a lot about how kids would learn a certain topic, a certain physics topic. For instance, he did some model cars with springs and we could argue about what the kids…not so much argue but he’d present what he wanted to and because I was learning the topic, we could take things apart and he cared enough to help me learn it along the way. And that happened with electricity as well. You know, I have some background on magnetism and electricity through the research stuff I’ve learned so I would pose an idea to him and he was say no I think it’s this way and I’d say no I think it’s this way. In order to present something to the kids we often had to work on things back and forth and I felt like I was learning more in-depth science from him but I was also formulating ideas that I’d been taught in theory, in education theory, and having to talk to him about that those were the things.”

“It [GK-12] has stretched me to branch out into topics that I am not an ‘expert’ in.”
“In order to present something to the kids we often had to work on things back and forth and I felt like I was learning more in-depth science from him but I was also formulating ideas that I’d been taught in theory, in education theory…”

Additionally, one fellow reflected:

“I am having a hard time recognizing that the teachers don’t have the level of expertise that I do in science and framing my interactions with them in a context relevant to their own experience and expertise. It can be hard to remember that in some ways, I need to act as an instructor to the teachers as well. Early on, I was feeling like I should be their student, which I still am to a certain extent, but also need to remember that we all learn from each other.”

In contrast, some teachers did not believe they learned anything new in math or science. One teacher explains,

“It’s not that I necessarily learned anything new, I just understand it better. I think [fellow] definitely contributed to that and I feel like I’d be able to get it to kids better.”

**Broadened Perceptions of Scientists**

At the beginning of the program, open-ended questions revealed that teachers were limited in their understanding of the fellows’ scientific research obligations. For example, one teacher explains:

“It took me some time for me to understand what they [the fellows] did outside of GK-12 and what the significance of their research was and it seemed like anytime they brought something to the classroom they had a much broader perspective and brought more worldly issues to the classroom, and I tend to be more local.”

At the end of the GK-12 program, it seems as though teachers had a better idea of fellows’ research and what they do as scientists. Several teachers said:

“I had no idea you could actually do some of the stuff that [the fellow] does…and the kids are learning about the different things they can do in their future.”

“…My students and I learned so much about the research process – the time it takes, why it’s important, the endless possibilities and opportunities available to my students. I also gained a new respect for the scientific process as seen through the eyes of a scientist.”
“[I have gained] a better knowledge of fellows and their work.”

An analysis of the teachers’ Draw-a-Scientist Checklist (DAST-C) revealed no significant differences for drawings before and after the GK-12 program. One reason for this might be that teachers drew a low number of stereotypes and a large number of positive attributes on the pre DAST-C, therefore there was not much room for improvement on the post DAST-C. Additionally, many teachers were already familiar with the DAST-C and had already assigned this activity to their students.

One possibility for the non-significant differences on the DAST-C in light of the focus group comments could be that teachers do not perceive “scientists” in a stereotypical manner even before their interaction with the GK-12 fellows. Although teachers may not have a stereotypical perception of a scientist, this does not mean that they are aware of what a scientist does and what a typical scientists’ research can encompass, as revealed in the focus group. It appeared that by the end of the GK-12 program, teachers had reached a broader understanding of the kinds of research conducted by scientists.

Discussion

This paper has described the Indiana Interdisciplinary GK-12 project and the evaluation model (logic model) used to measure progress toward desired outcomes as well as the results obtained from the first year’s data. Using a variety of evidence, results indicate that fellows, teachers, and students are meeting the desired goals of the GK-12 project.

Our evaluation of the GK-12 program contains features that are common to many educational program evaluations, such as low numbers of participants, a lack of random sampling, no comparable control groups, and time and budgetary constraints. In this paper, we show how the use of a mixed-methods approach to program evaluation can provide greater
interpretability of the results. A main strength of the mixed-methods evaluation design is the use of triangulation in situations where an experimental or even a quasi-experimental design may not be possible. In this case mixed-methods serve an important role in clarifying results and interpreting the program outcomes. As shown in the several of the program outcomes, quantitative data can provide a framework to help focus the qualitative inquiry. For example, our results for the fellows showed that a comparison of the fellows’ early concept maps to later representations revealed greater sophistication as evidenced by the use of cross-links. The concept maps were thus able to reveal gains when the quantitative STEBI-B could not, most likely due to the small sample size.

Further, when using small samples there will be reduced power, which makes significance more difficult to attain. Additionally, the time-consuming nature of qualitative methods such as focus groups may be balanced by the use of quantitative methods such as Likert-type surveys. The use of a mixed-methods approach may be necessary if drawbacks to conducting solely quantitative or qualitative assessments cannot be overcome.

One limitation of using a mixed-methods design for a large-scale evaluation project is the amount of data collected. Participants may feel that all of the data collection activities are time consuming. By far the most controversial evaluation tool has been the concept maps, especially for our participating teachers. Drawing and describing these maps was both time consuming and frustrating. It is important to emphasize the importance of the assessments not only for the reporting of project outcomes but also for the improvement of the program and individual participant’s outcomes. For the second year of the program we made additional efforts to do this.

One limitation that our program evaluation faced is the lack of available control groups of sufficient size. Although there were some significant differences on the student and teacher
surveys, without a control group it is not clear whether these differences were the result of the GK-12 program. We were able to obtain a control group for the fellows consisting of four doctoral students who had applied but were not admitted to the program. However, this number was too small to reveal any differences and the expense of providing compensation to this group outweighed the benefits. Consequently, the use of qualitative data, such as focus groups, interviews, and observations allowed us to better understand the possible impacts of the program.

An additional concern is the ceiling effect achieved on the Program Survey. The lack of variability in responses will create limitations in using this tool as a post-program measure of growth. Fortunately, by using a variety of both qualitative and quantitative measures we were able to capture some change. Triangulation using these diverse measures provided a more solid picture of program outcomes. Revision of the Program Survey instrument may be necessary for subsequent administrations.

The results prompted changes to the second year of the program evaluation. One of these changes not listed above is the use of an online survey format to minimize the time and expenditure for data-entry activities. Another change was to modify the Program Survey to include additional items to examine unexpected results that were discovered using the qualitative findings from the previous year. Therefore, the survey was modified for the second year to examine these changes quantitatively. Additionally, the DAST-C may not be appropriate for teachers who may already be familiar with this activity. Examination of the second year’s data may prompt an elimination of this activity for the teachers.

Carrying out evaluations on complex educational programs is a challenge but one that is necessary for reporting project outcomes, learning about a program’s effectiveness, and for program improvement. We hope that by providing an example of how we used mixed-methods
to evaluate the GK-12 project that those with evaluation needs for other educational programs might use our program as a guide. Instead of relying on quantitative methods alone, using mixed-methods to triangulate data and corroborate findings can lead to more confidence in the results and to a greater understanding of the program’s effects.

Another purpose of this paper was to provide results from our first year of the GK-12 program implementation. The results show that overall the GK-12 program appears to be meeting the desired goals for students, fellows, and teachers. Future directions will include strengthening our findings by analyzing the second year of program data.
References


National Science Foundation Graduate Teaching Fellows in K-12 Education (GK-12).

Retrieved on December 9, 2006 from:

http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=5472&from=fund


