Differences in Math and Science Understanding between NSF GK-12 Participant Groups: A Year Long Study

Jennifer Wilhelm
University of Kentucky

Xiaobo She
Texas Tech University

Darrellee Clem Morrison
Texas Tech University

Abstract

In this study, interdisciplinary environments were created in NSF institutes and classrooms with graduate fellows and teachers. Using a mixed methodology, we examined how experiential learning influenced participants' mathematical/scientific actions and compared differences in mathematics/science efficacy and content understanding between participant groups. Pre-assessment findings revealed no significant differences in content knowledge or self-efficacy issues between fellows and teachers. However, midyear and post test results showed a significant difference between fellows and teachers on the mathematics content domains. Significant differences were observed between gender groups (favoring males) on mathematical items of content and efficacy pre and midyear assessments and between science/mathematics discipline groups on mathematical items of content pre and post tests (favoring mathematicians). Interdisciplinary settings encouraged learners to voice their beliefs and realizations regarding their mathematical and scientific understandings. Most discomfort displayed by participants revolved around mathematics.

Objective

We describe interdisciplinary environments created in institute and classroom settings with 18 NSF GK-12 (National Science Foundation, Graduate and Kindergarten – grade 12) participants. The environments were designed with a framework that allowed participants the opportunity to experience the National Council of Teachers of Mathematics' (NCTM, 2000) and the National Science Education (National Research Council, 1996) standards in a blended and innovative manner. Through institute activities, participants investigated, analyzed, and communicated task findings. We compared differences in mathematics/science efficacy and content understanding between participant groups (i.e., gender, discipline, fellow/teacher status). We also examined how experiential interdisciplinary learning influenced participants' mathematical/scientific actions and beliefs.

Framework

Placing graduate students (fellows) and teachers together in a classroom environment has been done for nearly a decade through NSF GK-12 funding. The objectives are to improve communication and instructional skills for the fellows, and create content gains for teachers. Successes and challenges have been reported as a result of this program (deGrazia, Sullivan, Carlson, & Carlson, 2000; Genesan, Das, Edwards, & Okogbaa, 2004). For example, the program at the University of Colorado provided innovative ways for fellows to provide K-12 instruction in engineering technology. However, their evaluator reported that there was a constant need to "encourage teachers and fellows to schedule a regular meeting time to collaborate" (deGrazia et al., p. 21). The program at the University of South Florida had fellows developing curricula for teachers; however, "teachers felt unprepared…and would skip over the lessons and others would just touch upon the surface" (Genesan et al., p. 3).

Mitchell, Levine, Gonzalez, Bitter, Webb, & White (2003) stated in a GK-12 summary that one of the most claimed program accomplishments was that teachers "reported that their content knowledge had increased as a result of working with 'experts' in science" (p. 14). However, no evidence of this increase was provided except for the teachers' self-reports. In reviewing the literature on GK-12 programs, we found no studies that examined exactly what content knowledge was increased for participants in the program, whether teacher or fellow.

Research conducted by Schinske, Clayman, Busch, and Tanner (2008) documented ways in which the program enhanced children's self-efficacy of STEM content as they interacted with GK-12 participants. Schinske et al. reported high school students of GK-12 teachers engaging in analysis of scientific journal abstracts and figures. The purpose of these activities was to immerse students in authentic scientific activity. The post assessment data revealed that 40 of the 61 students claimed to have a more positive or more sophisticated view of scientific articles by the end of their school year compared to the beginning of these GK-12 activities. Once again, the data reported were only self-described testimonials and did not involve measurement of STEM content learned. Powers, Brydges, Turner, Gotham, Carroll, and Bohl (2008) conducted a study with children in schools involved in a Math Partnership that included GK-12 participants. The goal of their partnership was to positively impact the New York State mathematics exam scores. Their database of 3,748 students in grades 6-8 included a comparison group of 2,383 students who did not have participating GK-12 teachers and a treatment group of 1,365 whose math teachers received 60-90 hours of professional development.

> The students with a participating math teacher (treatment group = 67.94% proficient) outperformed students without a participating math teacher (comparison group=66.24 % proficient) in the grant. At certain grade levels the students of participating math teachers significantly benefited

> from their math teacher involvement in the grant. Students in the treatment group in grades five, six, seven, and eight especially benefited (Powers et al., p. 12).

The overall theme of the GK-12 program reported in this paper concerns the development of deep thinking about the integration of mathematics, science, and engineering by participating STEM research graduate fellows and secondary teachers. In this GK-12 program, we proposed that STEM fellows would have opportunities to 1) teach in a K-12 environment in a manner reflective of their increasing sophistication concerning the connection between math, science, and engineering as a tool for effective STEM education, 2) develop communication skills through dialogue with K-12 teachers regarding the ties between math and science in effective STEM education, and 3) produce through group teamwork (working with other fellows and teachers) an integrated math/engineering/science curriculum to improve K-12 STEM education.

The primary purpose of our program was to create opportunities for STEM graduate fellows to experience and develop their communication skills with fellow STEM researchers in other disciplines and within K-12 schools. This was intended to be accomplished through the formation of Mathematics/Engineering/Science Bridge (MESB) quartets (two fellows and two teachers) and through the quartets' development and enactment of integrated mathematics and science/engineering, inquiry-based curricular modules at the secondary level. All members of the quartet were required to meet on a regular basis to explore the relationships between mathematics and science/engineering fostered through the development of integrated modules. Due to additional in-kind funding from our university, we had the opportunity to add an extra engineering fellow and an extra engineering teacher to one group (making a sextet as opposed to a quartet). Therefore, during this first year in our GK-12 program, we ended up with three quartets and one sextet working within four area high schools.

Dr. Arden L. Bement, Jr., Director of the National Science Foundation, in testimony before the House of Representatives, has recently stressed the need for integrative STEM curricula (Bement, 2006). In a review of the efficacy of integrated STEM curricula, Wicklein and Shell (1996) determined that:

> After a careful examination of each of the pilot demonstration schools, three primary factors were identified that significantly affected the success or failure of the multidisciplinary curriculum: (1) teacher and administration commitment to the integration approach, (2) innovation and effort in curriculum re-design, (3) administration and teacher coordination of the integration plan. Each of these factors is of paramount importance in creating the type of integrated curriculum that will help students learn, apply, and transfer learning beyond the classroom environment.

"Connected knowing," which establishes a rich network of association, is among the most durable and accessible forms of knowledge. Connected knowing is most easily generated when "knowledge is primarily being constructed in interaction with other people, in a process that depends on understanding others' experiences, perspectives, and reasoning, and incorporates this understanding into the individual's knowing and understanding" (Boaler & Greeno, 2000, p. 174). Lave and Wenger (1991) and Adler (1998) might refer to this "connected knowing" as the theory of situated learning, which claims that learning and understanding is most conducive when interacting and participating with others. Situated learning transpires in authentic activities in which problem-solving strategies are used for real life advancements (Hendricks, 2001).

Lave and Wenger (1991) view schools that use traditional education programs as the institutions where knowledge is "decontextualized" and not transferable to real world working situations. In our GK-12 interdisciplinary endeavor, graduate fellows worked with other graduate fellows and with in-service teachers in which content knowledge and pedagogical content knowledge were shared to create a common discourse for understanding. The fellow and in-service teacher, the mathematics and science teacher, and the mathematics and science/engineering fellow relationships should result in outcomes that are greater than the sum of its parts, where participants learn new methods of developing knowledge and curriculum that can be transformed across boundaries, opening new perspectives for all learners.

We had hypothesized that strengthening the relationship among the mathematics, engineering, and science disciplines in both the educational and research venues would result in stronger, more cohesive learning and exploration in K-12 through graduate educational settings. It was proposed that integrated curricula was needed within the mathematics and science classrooms in order to allow learners the opportunity *to contextualize, to connect* to other disciplines, and to *experience* mathematical and scientific concepts.

Methods

This year-long study involved nine graduate student fellows (four from mathematics, five from science/engineering) and nine secondary teachers (four from mathematics, five from science/engineering). All fellows and teachers participated in a week-long summer institute and four half-day institutes during the fall and spring semesters. Table 1 displays the cohort information at each of the four area high schools.

During the summer institute, participants experienced interdisciplinary tasks enacted with varying degrees of inquiry. Participants were to consider the mathematics and science involved with understanding concepts within each activity. Tasks included: examination of population ecology using embedded statistics; modeling two-dimensional contour maps in three-dimensional space; exploration of variables that affect a pendulum's period; research on waves and superposition of sinusoidal functions; and investigations concerning relative age dating on a Martian surface using cratering rates and number density concepts. The purposes of these activities were to give participants opportunities 1) to experience interdisciplinary curricula as 'learners'; 2) to bond and communicate; and 3) to gain deep content connections and understandings.

During the fall and spring semesters, fellows were required to spend at least ten hours each week at their designated high school and up to five additional hours planning with their cooperating teachers. Both teachers and fellows were to keep weekly journal reflections documenting their experiences in the program and in the schools. Table 1 maps out the cohort arrangement and the percentage of time each fellow spent within the science/engineering classroom and the mathematics classroom. The fellows spent at least 50% of their classroom time in the environment of their own discipline. However, one fellow, Mr. L, did not have that option since he was an Electrical Engineer and the cooperating GK-12 science teacher was a Chemistry teacher. Another fellow, Ms. H, spent nearly 100% of her classroom hours with the Integrated Physics & Chemistry teacher instead of the Algebra teacher even though she was a Mathematics fellow.

Four half-day institutes occurred during the fall 2008 and spring 2009 terms. The purposes of the fall institutes were to give fellows the op-

Cohort School	Science/Eng. Fellow	Time %		Math Fellow	Time %		Science/ Eng. Teacher	Math Teacher
		Sci.	Math		Sci.	Math		
1	Mr. L (*Electrical Engineering)*	50	50	Mr. W	50	50	Ms. S3 (*11^{th} grade Chemistry*)	Ms. L (*9^{th} grade Algebra*)
2	Ms. S2 (*Chemistry*)	50	50	Mr. V	50	50	Ms. Y (*11^{th} grade Chemistry*)	Ms. K (*10^{th} grade Geometry*)
3	Mr. C (*Physics*)	55	45	Ms. H	99	1	Ms. A (*9^{th} grade Integrated Physics & Chemistry*)	Mr. D (*9^{th} grade Algebra*)
4	Mr. B (*Electrical Engineering*)	80 Eng. 5 Chem.	15	Mr. M	5 Eng. 15 Chem.	80	Ms. P (*11^{th} grade Intro. Eng.*) Ms. C (*11^{th} grade Chemistry*)	Mr. J (*9^{th} grade Algebra*)
	Ms. S1 (*Biochemistry*)	10 Eng. 70 Chem.	20					

Table 1. Cohort Information and Percentage of Time Fellows Spent in each GK-12 Classroom

Research Question	Data collection/instrumentation
1. How does experiential learning in interdisciplinary environments influence GK-12 participants' mathematical/scientific actions and/or beliefs?	Institute observations Written narratives
2. What initial differences in mathematics/science efficacy and content understanding exist between GK-12 participant groups (i.e., gender, discipline, and fellow/teacher status) and how do these differences change as participants engage in the GK-12 program?	Self-efficacy survey Content test

Table 2. Questions and Methods

portunities to present their current research via PowerPoint to the GK-12 audience and to give the four cohorts from each high school time to work on their integrated curricular modules. The purpose of the spring institutes was to allow each cohort the opportunity to enact their integrated modules they designed with the other GK-12 cohorts.

The focus of this study was to determine in what ways fellows and teachers changed or were influenced as they participated for one year in NSF GK-12 experiential activities. We specifically explored the research questions shown in Table 2.

This study utilized a mixed methods design where data collection included self-efficacy survey and content test responses (pre, midyear, and post), written narratives, and videotaped and transcribed institute observations. A mixed methods approach was employed with both quantitative and qualitative data sources and data analysis methods. Multiple research methods used in tandem allowed for triangulation and strengthened the study in greater ways than could have been done with either the qualitative or quantitative research alone (Creswell & Plano Clark, 2007).

Three researchers were participant observers throughout the institutes and were responsible for videotaping, administering assessments, and creating an on-line platform for participants' narrative responses. The qualitative data sources, videotaped interactions, and written narratives were used to determine participants' understandings and beliefs about inquiry-oriented, interdisciplinary curriculum and instruction. Analysis of the qualitative data was conducted through the exploration of patterns and themes. We used the practice of 'category construction' explained by Merriam (2009) as "I see a category the same as a theme, a pattern, a finding, or an answer to a research question" (p. 178). This process allowed us to uncover recurring themes that emerged from the data.

In addition to this qualitative data, we documented participants' initial, midyear, and post mathematical and scientific understandings and self-efficacy via a state standardized content assessment and a science and mathematics self-efficacy survey. We used this quantitative data to compare initial, midyear, and post differences in efficacy and basic content knowledge between participant groups (i.e., gender, discipline, fellow/teacher status). The content exam was designed to assess understanding of concepts such as algebraic patterns, geometry, statistics, problem solving, inquiry, physical and life science, and Earth/space science. The content exam used was an assessment that all pre-service teachers must pass in order to teach in the state's public schools. This particular version was actually designed for pre-service math/science middle level teachers. We chose to use the middle level assessment instead of the secondary assessment because we felt that this level of content (both mathematics and science) was the most appropriate for the purpose of measuring the basic content knowledge and skills possessed by our program participants. Also, the middle level content exam combined equal numbers of mathematics questions as science questions. There was no secondary level equivalent to this type of combined (mathematics and science) assessment. The self-efficacy survey used to assess the subjects' beliefs about their science/mathematics ability was adapted from an instrument developed by Ibe and Deutscher (2003) (see appendix).

Data & Analysis

For the data and analysis, we focus on a) patterns and themes within institute interactions and written narrations that illustrate how interdisciplinary experiences affect participants' actions and beliefs about mathematics and science, and b) group (i.e., gender, discipline, and fellow/teacher) comparisons of self-efficacy survey and state content exam results (pre, midyear, and post tests).

Institute Interactions and Narratives

We highlight three institute episodes that help to illustrate how experiential learning opportunities in a situated, interdisciplinary environment caused GK-12 participants' interests to be peaked, frustrations to be vented, and realizations to be expressed. The first episode, *The Pendulum's Period*, occurred on the second day of the five-day summer institute, and the second and third episodes, *The Border Problem and Cratering Rates on the Moon and Mars*, took place on summer institute day four. In order to further demonstrate the effect each episode had on participants' understandings and beliefs, we follow with their written narrative reflections.

The Pendulum's Period

Some participants found their institute experiences uncomfortable and challenging. We describe an episode where participants were examining a pendulum's period during the second day of the week long summer institute. The presenter gave each group equal masses, but different lengths of string, causing groups to record different periods. When a whole class plot of *Period versus Length* was made on the board, discourse ensued regarding the function type that appeared. Some remarked that "It looks like a square root function," and the instructor asked "How could you be certain?" One fellow stated that one could plot "*Log of Period versus Log of Length*", and the slope should equal ½. Another fellow said to plot "*Period Squared versus Length*", which should linearize the data. This discussion made some participants (both teachers and fellows) uncomfortable with several stating their students would "never" understand this because they did not understand. One teacher said that his students were taught to place time on the x-axis and it would be unnatural for them to place time on the y-axis. More discussion followed of independent and dependent variables. Some participants began to grasp and appreciate the mathematics that surfaced during these discussions and stated so during the institute and within their on-line reflections; others used the on-line platform to further express their frustration.

> "I got really frustrated…the discussions at the end of the lessons would be way over my students' heads. I learned a lot of new things, but the way that they were discussed was usually over MY head." (Ms. L, Mathematics Teacher)

> "It (*pendulum*) would be a good exercise to develop hypothesis on why the pendulum has the same period regardless of mass or why the pendulum exhibits a 'square root' behavior. Even if the background was not completely there, I found the enhancement of analytical/scientific thinking through applied math a priceless reward of this exercise." (Mr. L, Engineering Fellow)

> "I found the pendulum exercise most useful with respect to teaching Algebra. The math involved in this activity was the plotting of data. The realization that there may have been a connection with respect to the length of the string and the amount of time it took also sparked an 'aha' moment…this can be used for a variety of different graphing exercises, explaining what gives a graph a unique appearance, and for explaining square roots and quadratics in Algebra I." (Mr. J, Mathematics Teacher)

The Border Problem

A turning point occurred on day four of the summer institute. During the morning of day four, GK-12 participants heard a short presentation entitled "*Doing Mathematics: The Inquiry Way*" conducted by the first author. Embedded within the presentation was a break where learners had the chance to experience an inquiry-oriented mathematics problem connected to a real life situation. The proposed task involved solving "*The Border Problem*" borrowed from the Third International Math and Science Videotape Classroom Study (Stigler, Gonzales, Kawanaka, Knoll, & Serrano, 1999). The problem began by setting the task: Two property owners, Ms. Y (Chemistry Teacher) and Mr. C (Physics Fellow), share a border. They both wish to make their crooked border straight. Please find a way to make their border straight while conserving their original amount of land (see Figure 1).

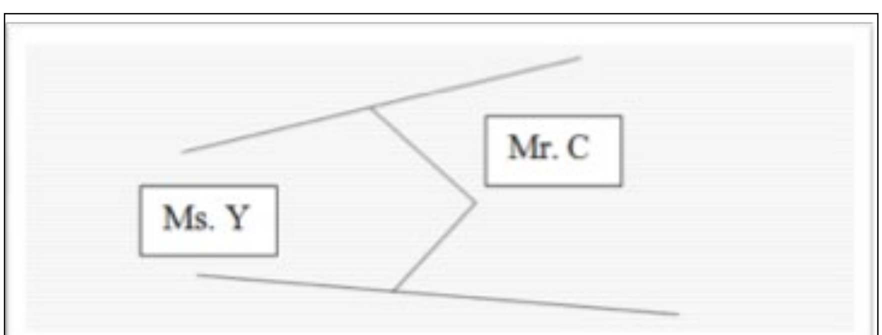

Figure 1. Border Problem Task with Left Side Property Owner Ms.Y and Right Side Property Owner Mr. C

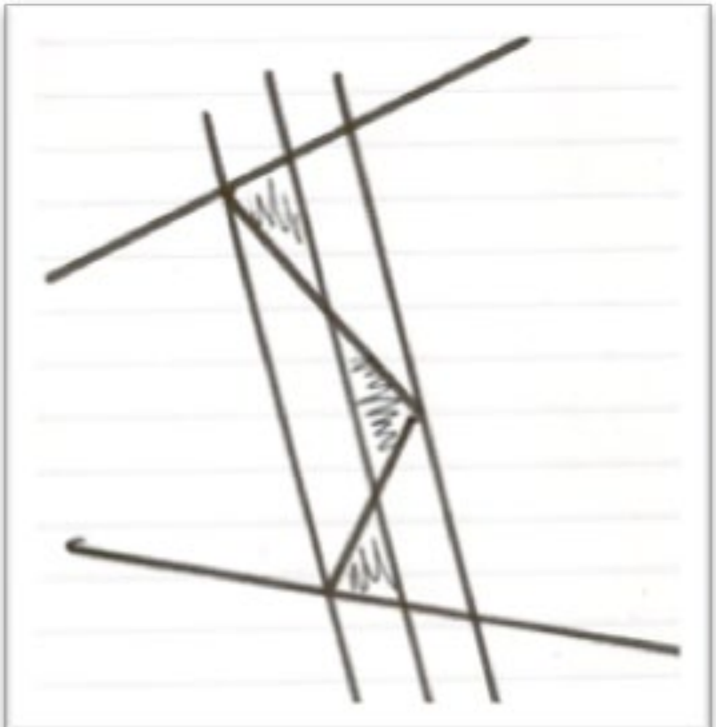

Figure 2. Representation of GK-12 Participants' Border Problem Solution

Learners were given ample time to solve the problem first working individually and then in their cohort groups. Each GK-12 group was then given the opportunity to present and communicate their solution. Most groups drew parallel lines and visually divided up the property equally, realizing their answers had room for error (see Figure 2).

Many found the problem difficult, and some stated that it would take extremely sophisticated methods to solve. Groups presented their solutions to the class. Following their solution presentations, a video excerpt was shown from the TIMSS Video Case Study where eighth grade Japanese students were solving the *Border Problem* in multiple ways using the idea that all triangles (with equal bases and equal heights) have the same area. No GK-12 participant solved the problem in this manner. While watching the video, many (both fellows and teachers) slapped their hands against their foreheads and exclaimed "Ahh!" to themselves.

At the end of this inquiry activity, people began to open up and express themselves, more than they had at any other time during the institute. Mr. J (Mathematics Teacher) later wrote in his reflections that "*inquiry-based teaching and learning can lead to outstanding results… the presentation also broke down some communication walls and people started to express their disagreements and opinions throughout the rest of the day.*" Fellows and teachers further communicated their thoughts in their online narratives regarding their *Border Problem* inquiry experience:

> "What a shocker, the problem was way easier than I thought." (Mr. L, Engineering Fellow)
>
> "(*The Border Problem*) module presentation was great. It was so funny to me to sit back and let the faculty and grad students grapple with such a simply solved problem. When we watched the video, I was amazed at how to solve the problem myself." (Ms. L, Mathematics Teacher)
>
> "I enjoyed the challenge of solving a problem that came from what I would call a scenario. Observing us as a combination of educators and graduate student experts, I was impressed to see the opportunities for each person to function at their own level of understanding in math." (Ms. A, Chemistry Teacher)

Cratering Rates on the Moon and Mars

Also, on day four (following *The Border Problem*), another GK-12 investigator presented an integrated science and mathematics scenario where the intentions were for participants to "be scientists." Learners were to examine features on the Moon and were asked to use their detective skills to identify four craters that could clearly be ordered from youngest to oldest. All were asked to explain their reasoning. This examination led into a discussion about the mathematical concept of superposition, and continued with additional investigation of cratering rates on the Moon and Mars. The final intended destination was for participants to investigate relative age dating on a Martian surface using these cratering rates. However, in the middle of their explorations, one GK-12 fellow, Ms. S2, voiced her concern regarding the direction and purpose of this astronomy-oriented activity. We follow with excerpts of the transcription:

> Ms. S2: Okay, going back to Earth now. This is a module, we're going to teach in the classroom? Right? What is the purpose of this?
>
> Researcher 1: Not necessarily. No, the purpose of this is to give you examples of the integration of math and science, and it's not necessarily something that you would use in your own classroom. You are learners here; in this workshop, you are a bunch of learners in the workshop. Whether or not you can or cannot use it in the classroom is really beside the point. It's this idea of you experiencing integrated math and science so that you can understand what it might look like.
>
> Ms. S1: I can use this in the classroom, basically…you don't really give the kids information, but you let them think about ages and…tell stuff is going to stack upon each other and you just tell, if you know that stuff stacks up you know, and the high stuff is the youngest, what kind of conclusions can you draw about the geography of the moon? You know, you don't necessarily have to know exactly what is going on, but you can draw some conclusions related to it. You know, I mean that's what we're doing right now.

This discussion continued on for several minutes of sometimes heated debate. It then took a turn towards how inquiry can or cannot be done in the high schools. One GK-12 participant stated that "*we have this ability to, you know, stay on task. This* (cratering rate activity) *would've been cut off a long time ago in a high school classroom*." More discussion ensued regarding teachers' having high and low student expectations. One Chemistry teacher complained about her experience trying to conduct inquiry in her classroom:

> "Those of us who have done some inquiry in the past several years, we didn't just start teaching that way…and if we started teaching that way, we were criticized because our classrooms were noisy, the kids weren't learning because they weren't in neat rows and weren't quiet…because our kids didn't learn to do that from the beginning…it's unnatural for our kids."

Despite some of the concerns displayed during this cratering rate lesson, some participants described in their on-line narratives their desire to learn more about these integrated concepts and their disappointment that the activity was not completed. Reflections were recorded on-line.

> "This module made me feel very uncomfortable since I had absolutely no knowledge of astronomy…I was completely blank and thrown out of my comfort zone. This experience was an 'eye opener' to understand the student in the classroom when they are asked to do something that they have not seen before. I really liked the module. Unfortunately, one student (*fellow*) raised the question 'where we were going with all this' instead of just learning from the experience; the student felt the need to just know the answer. I personally, believe this act was motivated because one was put out of the comfort zone…it was a very productive module that made me realize that I should give value to any participation the student have when investigating science." (Mr. L, Engineering Fellow)
>
> "I loved how (*Mars presenter*) started the problem…I was really able to evaluate my own scientific thinking, and by the end of the project, I felt I WAS A SCIENTIST! This method made some of the people in the room uncomfortable, and some questions were raised." (Ms. L, Mathematics teacher)
>
> "A GREAT model for Inquiry…I enjoyed participating in the lesson, but was frustrated at the end because I wanted to know more—much more. I was wishing I could leave, come home, grab my computer and start studying and researching. Wish I thought I could leave my own students with those same feelings." (Ms. A, Chemistry Teacher)

Three themes occurred throughout the written narratives and reflections of participants' institute experiences: 1) Interest and analysis of the summer institute's integrated mathematics and science modules; 2) Frustration when content, especially mathematical content, was or was not developed and/or thought to be presented in a manner "above the head" of participants; and 3) Expressions concerning communications about inquiry-oriented problem solving and realizations such as 'anyone can be a scientist' when given the opportunity. By the conclusion of the summer institute, it was clear that both fellows and teachers would be challenged as they negotiated their following semester commitments of designing and teaching lessons with an integrated, inquiry approach. We follow with quantitative results of groups' pre, midyear, and post efficacy and content understandings.

Self-Efficacy Survey and Content Exam

The content and science and mathematics self-efficacy assessments were given to the GK-12 participants prior to commencement of GK-12 activities (July 2008), midyear (January 2009), and as a post test (June 2009). A one-way analysis of variance (ANOVA) was conducted (pre, midyear, and post) on content and self-efficacy mean scores to determine if there were significant differences between groups (i.e., gender, discipline, fellow/teacher status). A repeated measures ANOVA was also conducted with the factor being gender, math/science, or fellow/teacher, and the dependent variables being the pre and post scores. This was conducted for each domain within each assessment as well as for the overall scores of each assessment.

The self-efficacy survey was administered to participants at the beginning of the institute and prior to the administration of the content exam. Reliability of both assessments was calculated using the *Cronbach's alpha*; this function measures the instrument's internal consistency. For the self-efficacy survey, the coefficient alpha was calculated for 0.87, and for the content exam, the coefficient alpha was calculated for 0.85. Both values were high and acceptable. The self-efficacy survey and the content assessment were re-administered midyear and at the conclusion of the program to capture midyear and final efficacy and content understanding. The initial test results aided the formation of strategic cohort groupings (balanced gender, content understanding, and discipline).

Self-Efficacy Results

Table 3 displays the statistics of the *self-efficacy survey* results by group. The survey used a Likert scale where the highest score of five shows the highest agreement with statements

	Gender			Discipline			Participant		
Survey	Males N=8 (SD)	Females N=10 (SD)	p-value	Science/Eng. N=10 (SD)	Math N=8 (SD)	p-value	Fellows N=9 (SD)	Teachers N=9 (SD)	p-value
Pre All	4.69 (0.24)	4.11 (0.49)	<0.01**	4.38 (0.40)	4.35 (0.61)	0.90	4.57 (0.34)	4.16 (0.54)	0.07
Mid All	4.48 (0.42)	3.76 (0.59)	0.06	4.18 (0.35)	3.96 (1.20)	0.58	4.08 (1.12)	4.08 (0.41)	1.00
Post All	4.45 (0.40)	4.07 (0.29)	0.03*	4.18 (0.32)	4.31 (0.47)	0.49	4.36 (0.31)	4.11 (0.43)	0.18
Pre Science	4.63 (0.34)	4.26 (1.00)	0.34	4.74 (0.26)	4.03 (1.03)	0.05	4.68 (0.39)	4.17 (1.00)	0.17
Mid Science	4.41 (0.53)	4.00 (1.08)	0.35	4.55 (0.28)	3.72 (1.16)	0.04*	4.18 (1.09)	4.18 (0.68)	1.00
Post Science	4.41 (0.56)	4.33 (0.58)	0.77	4.61 (0.24)	4.05 (0.69)	0.03*	4.49 (0.39)	4.24 (0.69)	0.36
Pre Math	4.70 (0.27)	3.76 (0.59)	<0.01**	3.91 (0.72)	4.52 (0.43)	0.05	4.38 (0.60)	3.99 (0.71)	0.23
Mid Math	4.47 (0.46)	3.36 (1.00)	0.01*	3.65 (0.67)	4.11 (1.25)	0.33	3.88 (1.19)	3.83 (0.77)	0.93
Post Math	4.44 (0.35)	3.61 (0.72)	<0.01**	3.58 (0.66)	4.48 (0.36)	<0.01**	4.13 (0.60)	3.83 (0.81)	0.40

Significance *p < 0.05; ** p < 0.01; ***p < 0.001

Table 3. Results of Self-Efficacy Survey by Groups

such as "I like science." The higher the score on the survey indicated a higher group efficacy concerning science and/or mathematics ability (see appendix). Significant differences between groups' mean scores were tested at the α = 0.05 level for the overall scores, the science scores, and the mathematics scores using a one-way ANOVA.

We found significant differences between male and female groups on the overall survey as well as on the mathematics items where males' self-efficacy scores were higher (for the pre, midyear, and the post tests). Significant differences were also found between science/engineering and mathematics participants on the science self-efficacy items of the midyear and post survey in favor of the science/engineers. Similarly, significant differences were observed between the discipline groups on the mathematics portion of the post survey in favor of mathematicians. No significant differences of efficacy were found between the fellows and teachers for the pre, midyear, or post survey.

Nearly every group's self-efficacy mean scores in science and mathematics dropped from pre to midyear, but then recovered slightly by the time of the post test. The only significant difference (over time) between mean scores occurred with the science participants on the science self-efficacy results. A repeated measures ANOVA for the science group showed a significant decrease from pre to midyear, $F(1, 9) = 7.642$, $p = 0.02$. A repeated measures ANOVA was also conducted with the factor being group (gender, discipline, or fellow/teacher participant) and the dependent variables being the pre and midyear test scores. This was conducted for both the science and mathematics domains of the self-efficacy assessment. No significant differences on interaction effect between any group and time were observed.

Content Assessment Results

Table 4 displays the statistics of the content assessment results by group. Significant differences between groups' mean scores were tested at the α = 0.05 level for the overall, the science, and the mathematics scores using a one-way ANOVA. Similar to the self-efficacy results, significant differences in mean scores were found between gender groups (favoring males) on the mathematics items for both the pre and midyear content assessment. However, by the time of the post assessment, significant difference between gender groups was no longer observed.

Significant differences in pre math mean scores between discipline groups (favoring mathematicians) were also observed. No other significant differences in pre mean scores were found between groups, including fellows and teachers. However, by the time of the midyear test there were significant differences in mean scores between fellows and teachers (favoring

	Gender			Discipline			Participant		
Content Test	Males %Correct N=8 (SD)	Females %Correct N=10 (SD)	p-value	Science/Eng. %Correct N=10 (SD)	Math %Correct N=8 (SD)	p-value	Fellows %Correct N=9 (SD)	Teachers %Correct N=9 (SD)	p-value
Pre All	74.9 (11.4)	66.2 (9.1)	0.09	67.8 (11.8)	72.9 (9.3)	0.34	73.9 (12.5)	66.0 (7.6)	0.14
Mid All	74.1 (17.3)	63.4 (10.5)	0.12	68.2 (11.3)	68.1 (18.7)	1.00	75.3 (10.3)	61.0 (15.1)	0.03*
Post All	74.5 (17.4)	65.6 (10.6)	0.20	67.8 (11.7)	71.8 (17.5)	0.58	77.0 (10.0)	62.1 (14.5)	0.02*
Pre Science	70.3 (13.0)	68.2 (13.6)	0.75	72.4 (12.1)	65.0 (13.7)	0.24	71.3 (14.0)	66.9 (12.3)	0.49
Mid Science	71.3 (17.9)	66.4 (19.8)	0.60	74.4 (13.5)	61.3 (22.3)	0.14	73.8 (13.5)	63.3 (22.1)	0.25
Post Science	70.0 (17.0)	67.2 (18.8)	0.75	72.4 (14.7)	63.5 (20.5)	0.30	74.7 (11.5)	62.2 (20.8)	0.14
Pre Math	79.5 (10.0)	64.2 (14.9)	0.03*	63.2 (15.2)	80.75 (6.8)	<0.01** (0.008)	76.4 (14.7)	65.6 (13.7)	0.12
Mid Math	77.0 (17.1)	60.4 (12.9)	0.03*	62.0 (14.9)	75.0 (17.0)	0.10	76.9 (11.0)	58.7 (17.1)	0.02*
Post Math	79.00 (18.5)	64.0 (14.0)	0.07	63.2 (13.5)	80.0 (18.1)	0.04*	79.3 (12.2)	62.0 (18.1)	0.03*

Significance *p < 0.05; ** p < 0.01

Table 4. Statistical Results of the Content Exam by Group (gender, discipline, fellow/teacher status)

fellows) on both the overall mean scores and the mean scores of the mathematics test items. These significant differences between fellow and teacher groups were also observed with the final post test results.

A repeated measures ANOVA was also conducted with the factor being group (gender, discipline, or fellow/teacher participant) and the dependent variables being the pre and midyear test scores or the pre and post test scores. This was conducted for both the science and mathematics domains of the content assessment. No significant differences on interaction effect between any group and time were observed.

Since the GK-12 program's objectives are to not only improve fellows' communication skills, but also to improve fellows' instructional skills, we analyzed the mathematics and the science content domains that dealt with learning, instruction, and assessment. Content domain six had eight test items concerning mathematical learning, instruction, and assessment; and content domain eleven had five items dealing with science learning, instruction, and assessment. No significant differences from pre to midyear or pre to post were observed in fellows' science domain eleven mean scores. However, a repeated measures ANOVA showed a significant decrease for the fellows' pre to midyear math domain six mean scores, $F(1,8) = 6.00$, $p = 0.04$, partial $\eta^2 = 0.429$. The partial η^2 of 0.429 indicates that approximately 42.9% of this score decrease can be directly attributed to the fellows' participation in the GK-12 program (summer and fall terms). By the time of the post assessment, fellows' mean scores on this domain increased and no longer resulted in a significant difference between pre and final post scores. Figure 3 displays fellows' and teachers' pre, midyear, and post mathematics learning, instruction, and assessment mean scores (domain six).

Two of the eight questions (of domain six) that had the greatest change from the correct pre to incorrect midyear answers (by fellows) were both test items dealing with student assessment. One of these questions described a scenario where a mathematics teacher used a variety of assessment methods, including multiple choice questions, constructed-response assignments, individual and group projects, peer- and self-assessment, student journals, and student portfolios. The question continued by asking what was the most likely benefit of this multiple assessment strategy. Six fellows that selected the correct response of "*allows students a variety of ways to show what they know and can do*" on the pretest, chose an incorrect response on the midyear test. Five of these six people instead selected "*allows the teacher to eliminate potential bias in assigning students' grades.*" The reason for this change in answer choice could possibly be due to the fact that on the last day of the summer institute, a mathematics specialist from the local school district conducted a presentation regarding state standardized testing and problems with question bias. In depth discussion followed concerning difficulties with state standardized testing (including test bias) and the burden this testing places on public school teachers.

Another possibility could be the empathy generated in fellows as they experienced their first semester in the high school classroom. Fellows' eyes were opened to the lives many of the high school students led and to the apparent inequities and bias in the public schooling system. The following two narratives written at the beginning of the fellows' second semester in the schools are representative of this empathy.

"Today was an interesting day in Chemistry. One of Ms. S3's entire classes got moved out by administration to make room for an AP Chemistry class…I don't know, something about it just seems wrong. Of course, I'm not lying when I say that most of the students probably enjoyed the move! The new AP class was so much smoother though…I had never seen so many white kids together in a classroom at that school until that class

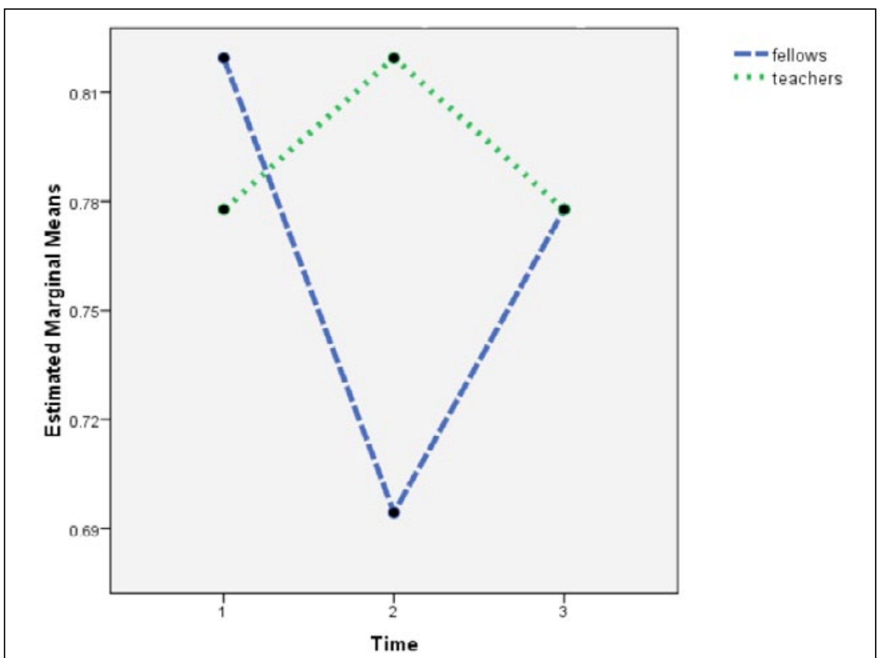

Figure 3. Pre (1), Midyear (2), and Post (3) Mean Math Domain Six Scores by Fellows and Teachers

came in. I don't know… (*is it*) unintentional (I hope) discrimination among the administration. I think I'd like to see a hispanic teacher teach a mostly hispanic class; I wonder if the students would be more apt to listen; most of the teachers I've seen are white." (Mr. W, Mathematics Fellow)

"I am mostly extremely surprised at how unprepared I am for this. Also, I feel for these kids. There are many days that I leave in tears because of them. I hate how some of them have to live…I hate that they have no self-esteem. I hate that you can almost literally see them falling through the cracks. And I don't know what to do. I work one-on-one with them as often as possible and try to encourage them about there being life after high school and childhood…I don't think I can do this for much longer; I don't know how teachers do it…I don't know what on earth I can teach them that will really benefit them and make the bad stuff stop." (Ms. J, Mathematics Fellow)

By the time of the post test given in June (2009), fellows' scores increased on math domain test items (still lower than where they had begun at the beginning of the program). The teachers' scores decreased from midyear to post, landing them with the same average final score as the fellows'.

Discussion

Pre-assessment findings revealed no significant differences in content knowledge or self-efficacy issues between fellows and teachers, which is contrary to the NSF GK-12 premise regarding fellows as the content "experts" of the classroom. However, midyear and post test results showed a significant difference between fellows and teachers on the mathematics domains of the content assessment mainly due to the teachers' mean scores decreasing as the fellows' mean scores increased. In addition to the differences between fellows and teachers, we found significant differences between males and females (favoring males) on the mathematical items of both content and efficacy pre and midyear assessments. We also discovered significant differences between science/mathematics discipline groups on the mathematics portion of the content pre and post tests (favoring mathematicians).

The quantitative data results revealed that the underlying difficulties and source of greatest concern with content had to do mainly with mathematics, which was also observed with the qualitative data outcomes. Most of the discomfort displayed in institute interactions and within written narratives revolved around mathematics. In fact, right from the commencement of the GK-12 activities, both the quantitative pre-assessments and the qualitative institute interactions displayed some participants' reluctance to engage in mathematics in relevant, meaningful, or deep ways. However, we found that a situated, interdisciplinary setting during the summer institute initiated learners to voice their concerns, beliefs, and realizations regarding their own mathematical and scientific understandings as well as secondary students' abilities. Although such realizations were made early on, there was no quantitative evidence of significant growth in the content arena.

Two significant decreases from pre to midyear assessment mean scores were observed: 1) Between pre and midyear mean science self-efficacy scores of the GK-12 science/engineering participants, and 2) Between the pre and midyear domain six (math learning, instruction, and assessment) content mean scores of the GK-12 fellows. It is unclear to what the causes were for the significant decline on these assessment items of both the content exam and the self-efficacy survey. However, there does appear to be evidence of inconsistencies between what the NSF GK-12 program purports to accomplish and the actuality of the everyday classroom experienced by students, teachers, and fellows. Both significant decreases disappeared by the time of the final post test, which illustrates that the midyear snapshots were instrumental towards capturing midyear difficulties and possible program deficiencies.

Importance

Collection and analysis of this data was vital to ensure strategic grouping (males and females, science/engineering and mathematicians, fellows and teachers) of GK-12 participants in the schools and to aid the design of institutes throughout the academic school year. This data also served as three benchmarks of participants' initial, midyear, and post perceptual, conceptual, and pedagogical understandings of STEM disciplines, which may be the first evidence (besides self-reports) in the history of the NSF GK-12 program. The midyear data was especially useful to gauge the fluctuations that occur midway through a year-long program and to inform decisions for the remainder of the academic year.

Due to the limitation of the sample size, this study cannot be generalized to represent all

NSF GK-12 participants and programs. However, by an in-depth mixed methods analysis, we were able to shed light on potential problems and concerns regarding NSF GK-12 programs. Future studies should measure the development of GK-12 participants' content learning across multiple programs.

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. 0742402.

References

1. Adler, J. (1998). Lights and limits: Recontextualising Lave and Wenger to theorise knowledge of teaching and of learning school mathematics. In Watson, A. (Ed.) *Situated cognition and the learning of mathematics.* Centre for Mathematics Education Research. University of Oxford, Department of Educational Studies, Oxford, 161-177.

2. Bement, A. (2006) Testimony before the committee on science, *U.S. House of Representatives Hearing on K-12 Science and Math Education across Federal Agencies*. Retrieved on February 20, 2009 http://www.nsf.gov/about/congress/109/alb_k12edu033006.jsp

3. Boaler, J. and Greeno, J. (2000). Identity, agency, and knowing in mathematical worlds. In J. Boaler (Ed.), *Multiple Perspectives on Mathematics Teaching and Learning.* Westport, CT: Ablex Publishing, 171-200.

4. Creswell, J. W. & Plano Clark, V. L. (2007). *Designing and conducting mixed methods research*. Thousand Oaks, CA: Sage.

5. deGrazia, J. Sullivan, J. Carlson, L. & Carlson, D. (2000). Engineering in the K-12 classroom: A partnership that works. *Proceedings of the American Society for Engineering Education/ Institute of Electrical and Electronics Engineers (ASEE/IEEE) Frontiers in Education Conference*, (Kansas City, MO), p. 18-22.

6. Genesan, R. Das, T. Edwards, C. & Okogbaa, G. (2004). Challenges in enhancing science education in elementary classrooms through university-school district partnerships. *Proceedings of the American Society for Engineering Education/ Institute of Electrical and Electronics Engineers (ASEE/IEEE) Frontiers in Education Conference*, (Savannah, GA), p. 1-5.

7. Hendricks, C. (2001). Teaching casual reasoning through cognitive apprenticeship: What are results from situated learning? *The Journal of Educational Research,* 94(5), 302-311.

8. Ibe, M. and Deutscher, R. (2003). The role of the goldstone apple valley radio telescope project. Paper presented at the annual meeting of the American Education Research Association. (Chicago, Illinois).

9. Lave, J. And Wenger, E. (1991). *Situated Learning: Legitimate Peripheral Participation.* Cambridge, United Kingdom: Cambridge University Press.

10. Merriam, S. B. (2009). *Qualitative research: A guide to design and implementation*. San Francisco, CA: Jossey-Bass.

11. Mitchell, J. Levine, R., Gonzalez, R., Bitter, C., Webb, N. & White, P. (2003). Evaluation of the National Science Foundation Graduate Teaching Fellows in K-12 Education (GK-12) Program. *Paper presented at the Annual Conference of the American Education Research Association*. (Chicago, Illinois).

12. National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, Va.: NCTM.

13. National Research Council. (1996). *National science education standards*. Washington, DC: National Academy Press.

14. Powers, S.E., B. Brydges, P. Turner, G. Gotham, J.J. Carroll, D.G. Bohl Successful Institutionalization of a K-12 - University STEM Partnership Program. *In: Proceedings of the 115th Annual ASEE Conference & Exposition* (Pittsburgh PA, June, 2008, on CD, Session # AC 2008-1652)

15. Schinske, J., Clayman, K., Busch, A., and Tanner, K. (2008). Teaching the anatomy of a scientific journal article: three activities engage students in the process of science. *The Science Teacher*, 49-56.

16. Stigler, J.W., Gonzales, P., Kawanaka, T., Knoll, S., and Serrano, A. (1999). *The TIMSS Videotape Classroom Study: Methods and Findings from an Exploratory Research Project on Eighth-Grade Mathematics Instruction in Germany, Japan, and the United States. (NCES 1999-074). U.S. Department of Education*. Washington, DC: National Center for Education Statistics.

17. Wicklein, R. and Schell, J.(1995). Case studies of multidisciplinary approaches to integrating mathematics, science and technology education. Journal of Technology Education 6(2). Retrieved on February 20, 2009 http://scholar.lib.vt.edu/ejournals/JTE/v6n2/wicklein.jte-v6n2.html

Dr. Jennifer Wilhelm joined the University of Kentucky faculty in 2009 as an Associate Professor in Science Education and as an engagement and outreach faculty member of the UK Partnership Institute for Mathematics and Science Education Reform (PIMSER). She holds an M.S. in Physics from Michigan State University and a Ph.D. in Mathematics/Science Education from the University of Texas at Austin. Dr. Wilhelm's primary research interest involves the design of interdisciplinary, project-based learning environments. She investigates how people understand science and mathematics concepts as they participate in project work that demands the integration of multiple content areas.

Xiaobo She is a doctoral candidate in the College of Education at Texas Tech University, majoring in Curriculum and Instruction with an emphasis in mathematics education at middle levels. Recently, she has been involved in several research projects, working with both pre-service and in-service teachers in the area of mathematics. Her primary research interest focuses on comparative studies across countries and intervention design to improve mathematical competency among pre- and in-service teachers. Ms. She has been certified as a middle school and high school mathematics teacher.

Darrellee Clem Morrison is a doctoral candidate at Texas Tech University in the College of Education. She holds an M.S. in Curriculum and Instruction from Texas Tech University and has been conducting research in the areas of mathematics and science education at the middle and secondary levels for the past three years. Mrs. Morrison plans on teaching mathematics at the secondary level upon completion of her dissertation focusing on post secondary math and science.

Appendix: Science and Mathematics Self-Efficacy Survey

Gender: male female Name:

Major (level): Date:

Scientific and Mathematics Efficacy Instrument Adapted from:

Ibe, M. and Deutscher, R. (2003). The role of the goldstone apple valley radio telescope project. *Paper presented at the annual meeting of the American Education Research Association*. (Chicago, Illinois).

Please select and record the scale number that sounds most like you for each of the following items.

1(Strongly Disagree), 2(Somewhat Disagree), 3(Undecided), 4(Somewhat Agree), 5(Strongly Agree)

1. I like science _____
2. I can think scientifically _____
3. I am able to use scientific equipment _____
4. I can be trusted to communicate scientific concepts accurately _____
5. I am comfortable asking scientific questions _____
6. I value my work as a scientist _____
7. Other people value my work as a scientist _____
8. I am able to find mathematical patterns or trends in my data _____
9. I like mathematics _____
10. I can think mathematically _____
11. I am able to use mathematical software programs _____
12. I can be trusted to communicate mathematical concepts accurately _____
13. I am comfortable asking mathematical questions _____
14. I value my work as a mathematician _____
15. Other people value my work as a mathematician _____
16. I am able to find possible scientific reasons to explain patterns in data _____
17. I can rethink my ideas based on new information _____
18. I can complete a long-term project _____