Learning about Space Science: Comparing the efficacy of reform based teaching with a traditional/verifications approach

E.M. Granger and T.H. Bevis,
Office of Science Teaching Activities
College of Arts & Sciences
Florida State University

Y. Saka and S.A. Southerland
School of Teacher Education
Florida State University

ABSTRACT

This research explores the relationship between curriculum and teachers' knowledge and beliefs about teaching. Using a quasi-experimental design, the effectiveness of Great Exploration in Math and Science (GEMS) Space Science Curriculum Sequence was compared with the effectiveness of more traditional curriculum in supporting 4th and 5th grade teachers' learning of space science as well as their knowledge and beliefs associated with the teaching of science. *GEMS* employs an inductive approach to content (learning cycle), explicit use of evidence, and attention to scientific inquiry. Randomization occurred at the level of the teacher assignment to treatment group (not at the student level). The sample included 32 treatment and 29 control teachers. Our findings suggest that reform-based curricula combined with professional development around the curricula can be effective in shaping teachers' content knowledge and beliefs about teaching. The GEMS materials were more effective in supporting the professional development of teachers who had more to learn (i.e., teachers with lower self efficacy and teachers with lower outcome expectancy at the outset), and the effect of the use of GEMS lessened for teachers who had high self efficacy and outcome expectancy at the outset of the study.

Corresponding author
E.M. Granger
granger@bio.fsu.edu

Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA, April 12-17, 2009.

Learning about Space Science: Comparing the efficacy of reform based teaching with a traditional/verifications approach

Introduction

Increasingly there are calls for science classroom practice to be situated in the central tenets of reform-based teaching practices (e.g., AAAS, 1996; Duschl, Schweingruber, & Shouse, 2007; Glenn Commission, 2000; National Academy of Sciences, 2007; NRC, 1996; National Science Board, 2006). Recent studies on learning science emphasize three key areas critical to student success: (1) understanding the discipline's core concepts, theories, and models; (2) understanding, in general, how scientific knowledge is generated, tested, and accepted; and (3) using the first two to extend understanding into new areas. Research suggests that in order for students to come to understand scientific concepts, be able to apply the concepts, and understand how science is done, their learning of science should in some ways echo the way science is conducted by scientists (e.g., Bybee, 1997; Chinn et al., 2002; Duschl et al., 2007; Flick, 2003).

The findings of science education research and the central features of the national reform efforts have been applied to produce promising, well-designed instructional materials that have undergone years of development, field-testing, and revision (e.g., *Great Explorations in Math and Science*, *Insights and Outcomes, Full Option Science Systems, Math Connections, Connected Mathematics Project, BSCS*). However, despite extensive efforts to precipitate educational change using such materials, classroom practices remain largely unaffected. One of the reasons for this lack of use of these materials has to do with teachers' reticence to use these methods in the face of accountability pressures (Abrams Southerland, & Evans, 2007) as teachers remain unconvinced that reform-based practices (e.g., levels of inquiry, learning cycle, writing to learn) are successful in helping students learn science concepts (Settlage & Blanchard, 2007; Settlage & Meadows, 2002). Teacher reticence is also based, in part, on their discomfort with and/or lack of knowledge about the content itself (including the nature of science and the processes of science that produce the knowledge), the teaching practices recommended by current research, and the pedagogical content knowledge to enact both (content and practices) in the classroom (Yore et al., 2007). Reform-minded teaching places many demands on teachers, making it hard to implement without well-designed curricula *and* professional development to support teachers as they change their classroom practice (e.g., Blanchard, Southerland & Granger, 2009; Crawford, 2000, 2007; Luft, 2007; Windschitl, 2004).

One of the central findings of the analysis of past reform efforts is that teachers' personal characteristics are very influential in shaping their enactment of reform-based practices (Gess-Newsome et al., 2003). Such characteristics as teacher content knowledge, self efficacy, interest or attitude, views on inquiry learning, self confidence, self doubt, and outcome expectancy have been found to shape science teaching practice (e.g., Ashton & Webb, 1986; Bandura, 1977; Cronin-Jones, 1991; Krajcik et al., 1994;

Rice & Roychoudhury, 2003; Ross, 1992; Settlage et al., 2008; Tschannen-Moran et al., 1998; Woolfolk et al., 1990; Wheatley, 2002).

Crawford (2000) describes that teachers' content knowledge, pedagogical content knowledge, and theoretical knowledge must act together in order to enact the features of science education reform—a tall bill to fill given the day-to-day pressures of present day teaching. Further, this amalgam of teacher knowledge that Crawford describes must be enacted in the context of a classroom and school in order for the practices of reform to be realized (Gess-Newsome et al., 2003; Yore et al., 2007). One central feature of this context is the curricula teachers have available and are expected to employ. In our work with teachers, they often request reformed-based curricula that are readily implementable to help them progress along the continuum toward reform-based practices in the teaching of science (just as has been found productive in mathematics) (e.g., LaChance, 2004; Schneider et al., 2005).

The construct of teacher professional development hinges on concrete classroom applications of general ideas; it involves opportunities for actual practice rather than mere descriptions of practice. Although observation, critique, reflection, group support and collaboration, and evaluation and feedback from skilled practitioners (e.g., professional developers, university faculty, science coaches) do help if the practitioners have a wealth of experience in teaching, teacher development ultimately depends on opportunities for actual classroom practice (Elmore & Burney, 1997; Sadler, 2006). Throughout the literature, teachers' professional development has been understood as a key component of the implementation of educational reform (Southerland et al., 2007; Yager, 1992). However, the current reform efforts are undermined by unrealistic expectations from teachers who do have a wealth of classroom practice and find it difficult to change their practices. For example, McIntyre and Hagger (1992) state:

> Development takes what is there as a valuable starting point, not as something to be replaced, but a useful platform on which to build. To do so is to recognize not only that teachers do have valuable existing expertise but also that, if teachers are forced to choose, they will usually revert to their secure established ways of doing things. The metaphor of 'building on what is already there' is not, however, satisfactory because it suggests adding on something separate to what is there; something extra on top. The concept of development, in contrast, implies that whatever is added, whatever is new, will be integrated with what is there already, and will indeed grow from what is there. (p. 271)

At the center of the discussions about effective professional development opportunities for teachers, Feldman et al. (2009) proposed that there is a critical need for the science education community to provide such opportunities for science teachers to enhance their practices in a way that is mindful of science education reform. They further suggest revisiting existing professional development opportunities to refine their structure so as to meet practicing teachers' abilities and needs. We argue that one notion helpful in this regard is to help teachers enact reform-based practices through reform-oriented curricular materials in combination with professional development.

This study follows a host of others (Blanchard et al., 2009; Gess-Newsome et al., 2003; Smith, 2005; Smith & Southerland, 2007) that investigate the interaction of context and practice with teachers' knowledge and beliefs about science teaching. The facet of context that we emphasize is the combination of professional development and reform-based curricula. A goal of this research is to study, through a randomized, quasi-experimental design, the influence of both professional development and the enactment of reform-based science curricula on fourth- and fifth-grade science teachers' knowledge and beliefs about space science and the teaching and learning of science.

The new GEMS *Space Science Curriculum Sequence* (hereafter referred to as GEMS *SSCS*) for grades 3-5 (Lawrence Hall of Science, 2007) was chosen for this study because it embodies many of the qualities of the reform-based approach to science teaching as described in the *National Science Education Standards* (NRC, 1996), *Inquiry and the National Science Education Standards* (NRC, 2000), *Taking Science to School* (Duschl, et al., 2007), and *Ready, Set, Science* (NRC, 2008). Additionally, it exemplifies curricula designed to enable teachers to provide reform-minded instruction for their students, including extensive classroom lesson plans complete with detailed instructions for enactment, questioning and responding outlines, detailed content background appropriate for elementary-school teachers' understanding, pedagogical professional development notes for the teacher about the instruction (learning cycle, leading discussions, nature of science, etc.), timeframes, critical junctures, and formative and summative assessments. Furthermore, the space science content, science process, and nature of science content embodied in this curriculum are aligned with the national science standards and those of the state in which the study occurred.

We report here some preliminary results from the data collected during the first year (cohort one) of a two year (two cohort) study. This paper will focus on the effects of professional development and teaching with the GEMS curriculum compared to the effects of teaching with a more traditional curriculum on five teacher outcomes: content knowledge, confidence, self efficacy, outcome expectancy, and attitude. Overall, our work is designed to begin to statistically address several questions surrounding enactment of reform-based curriculum following professional development: Does the use of such curricula plus professional development work? How does it work, that is, what mediating variables can we identify that indirectly transmit the effect of the treatment? For whom does the use of such curricula plus professional development work, that is, what student moderator variables can we identify from the data collected? Under what conditions do such curricula plus professional development work, that is, in this case what teacher moderating variables can we identify from the data collected? This paper will report on some preliminary answers to the latter question, allowing us to begin to identify some of the characteristics that made teachers more effective in facilitating the desired student outcomes.

Methods

Overview

The portion of our study reported herein employed a randomized quasi-experimental design. The experimental treatment group was comprised of 32 teachers who participated in a 4 day professional development experience focusing on the GEMS *SSCS* and the pedagogy underpinning it (i.e., learning cycle, evidence circles, cooperative learning, discussion techniques, space science misconceptions, guided [level 1 and 2] inquiry) and who then enacted this curriculum in their classrooms. At the same time the control group was comprised of 29 teachers (3 were dropped from the study due to either personal issues or non-compliance with data collection procedures) who participated in more limited professional development and who used the district adopted text to address the same space science content, science process, and nature of science standards through traditional, transmissive-mode approaches to instruction (lecture, reading from the text, and "hands-on" activities that are related to the topic but do not extend the depth of the student learning about core concepts nor address misconceptions). Teacher knowledge of space science and affective dimensions related to science teaching were compared across groups.

Participants

This study was conducted in a county in central Florida in 4th and 5th grade classrooms during the 2007-08 school year. Randomization occurred at the level of the teacher assignment to treatment group (not at the student level). The year one sample originally targeted 36 experimental treatment group teachers and 36 control group teachers, but by the completion of year one there were 32 treatment and 29 control teachers remaining in the study (Table 1). Twenty of these teachers or 32.8% were fourth-grade (10 treatment/10 control) and the remainder was composed of fifth-grade (22 treatment/19 control) teachers. Their years of experience ranged from first year teachers to teachers with 38 years of experience, with the average being 10.9 years; the treatment group averaged 10.4 years and the control group averaged 11.4 years. All teachers in both groups were certified. Teacher volunteers were assigned to treatment or control group with control/experimental group matching according to grade level, SES, school grade, and ethnic diversity based on their students' demographics.

Table 1: Teacher Data

	Average Years of Experience	**Grade Level**	
		4th	**5th**
GEMS Treatment (n=32)	10.4219	10	22
Control (n=29)	11.3621	10	19

Curricula

The curriculum enacted by the experimental treatment group was the GEMS *Space Science Curriculum Sequence (2007),* which is a curriculum for teaching space science concepts for grades 3 through 5. Through experiential learning, discussions, and reflections centered upon the pedagogies that underpin the GEMS *SSCS* teachers were prepared for teaching using this very "teacher friendly" curriculum. We label this curriculum as "teacher friendly" in that it was easily implementable by teachers inexperienced with reform methodologies given its detailed materials preparation and classroom enactment instructions, including time frames for preparation and teaching, scripting of lessons and discussions, embedded formative and summative assessments, etc.. This GEMS curriculum was designed to address age-appropriate core concepts in space science (NSES, 1996) and common misconceptions that students harbor about them (Kavanagh, 2007a, b). The activities in the curriculum specifically target these core concepts and misconceptions in the attempt to change students' knowledge. In addition, the curriculum has an explicit focus on the role of models and evidence in science. In it students are encouraged throughout to evaluate alternative explanations, to use evidence to support explanations, and to critique the merits of an explanation in a scaffolded, age-appropriate way. In general, all lessons were structured around a learning cycle format (e.g., Bybee, 1997). Teachers in the experimental group were further instructed by the research team to adjust their normal classroom practice, to closely follow the instructions described in the curriculum.

The GEMS curriculum contrasts with the district adopted science text for grades 4 and 5 that served as the basis of the control group classroom instruction. The district curriculum was centered on more didactic presentation of space science concepts including direct instruction, reading of text, students answering very focused questions. The activities included in the text typically served as verification of the content already presented in the text or, more commonly, activities that were peripherally associated with the topic, but that did not address the core concepts. Control teachers were further instructed by the research team to adjust their normal classroom practice, if it was different from this model employed in the text, in order to follow text's presentation for the space science unit.

The structure of the experiment in which they were participating (i.e., quasi experimental design) was discussed with both groups of teachers. The importance of such studies to the teaching profession was stressed as was the importance of their contributions to their profession through participation in this study. In this context, teachers in both groups were instructed to refrain from adding any additional activities to those present in their assigned curriculum. Teachers from GEMS treatment and control groups in the same school (and this was done for a matched control whenever possible) were instructed not to discuss their curriculum with the other group until after the administration of the delayed post testing five months after the teaching of the unit. Teacher fidelity to the assigned methodology was assessed through direct observation and/or videotaped observation at least twice during the unit using the RTOP instrument (Sawada et al., 2002). (Analysis of all videotapes for cohort one is not yet complete at this time.)

Professional Development

Because we realize that teachers' use of any new curriculum is an act of interpretation, the GEMS treatment group was involved in professional development in which the teachers experienced the curriculum as learners, then learned about the pedagogies that underpin it (e.g., learning cycle approach to science instruction, questioning/discussion strategies, evidence circles, assessment strategies, nature of science teaching strategies, etc.) through an explicit/reflective experiential approach. This occurred in a 4-day professional development workshop two weeks before the beginning of school, a three-hour follow up immediately prior to the beginning of the teaching of the space science unit, and a three-hour session midway through the teaching of the unit to discuss questions that had arisen. A "science coach" was available to help them with logistical or pedagogical issues. In actuality the support requested turned out to be wholly logistical.

Teachers in the control group participated in a meeting 3 weeks prior to the teaching of the control space science unit to review their part in the project, to discuss the need for employing traditional teaching approaches through fidelity to the district-adopted textbook curriculum and the traditional approach to teaching asked of them (see above), and to complete the teacher pre-assessments.

Data collection

Five instruments were used to assess teacher learning, concepts, and affective dimensions (pre, post). These instruments were: Space Science Content test (Sadler et al., 2006), the Test of Science Related Attitudes (TOSRA) (Fraser, 1978), the Beliefs about Reformed Science Teaching and Learning assessment (BARSTL) (Sampson & Benton, 2006), the Teaching Science as Inquiry assessment (TSI, Dira-Smolleck, 2004), and the Views on Science Inquiry assessment (VOSI) (Lederman & Lederman, 2005). Semi-structured interviews using the Teacher Belief Inventory (e.g., Luft et al., 2003) were conducted with each teacher post teaching.

For the GEMS treatment group, each assessment was administered prior to professional development and approximately two weeks following completion of teaching the space science unit. For the control group the initial testing was completed 3 weeks prior to teaching of the textbook-based space science unit; post testing was accomplished on the same schedule as for the treatment group relative to teaching of the unit.

Data Analysis

The effects of the GEMS treatment on five teacher outcomes from three of the teacher assessments (Space Science Content, TSI, and TOSRA) are reported herein. These effects were obtained with standard single-level multiple regression models. To estimate the main effects of the GEMS treatment, the independent variables in the final models were the coded GEMS treatment variable (1 = GEMS group and 0 = Control group) and the teacher pre-measure corresponding to the outcome. The final models were the result of initially assuming models that allowed the GEMS effects to be different across the two grades (fourth and fifth). This was accomplished by including a coded grade variable (1 = Grade 4 and 0 = Grade 5) and the GEMS by grade product. In addition, the teacher years of experience (Experience) variable was also included in the initial models. The grade, grade by GEMS, and Experience variables were not significant in these initial models and were dropped for the final analyses. To test possible interactions of the GEMS treatment with the corresponding teacher pre-measure of the outcome (i.e., to test the pre-measure as a possible moderator of the GEMS effect), the appropriate product term was added to the main effect models.

Given the different scales of the outcomes of interest, standardized effects are included in the results in order to allow direct comparison of GEMS effect sizes for the various outcomes. These were obtained by dividing the estimated raw score GEMS effect by the standard deviation of the outcome variable. The strengths of the effects were characterized based on the following definitions for standardized effects (Cohen, 1977): 0.2 is "small," 0.5 is "medium," and 0.8 is "large." The results herein contain many statistical hypothesis tests, resulting in substantial inflation of family-wise error rate. Since this study is viewed as exploratory, there was no attempt to control family-wise error.

Results

The results reported herein will focus on some of the data collected from the teacher portion of this study. Nevertheless, it is important to situate these results in the context of the overall student outcomes of the study. Briefly, student achievement on the post content test (questions from Sadler et al., 2007) for GEMS group compared to the control group was positive and statistically significant ($p=0.004$). Student attitude (Homerton Science Attitudes survey, Warrington et al., 2000) for the GEMS group as compared with the control group was positive and statistically significant ($p=0.067$). Student delayed post testing (5 months ± 2 weeks) for each of these dimensions was in the positive direction, but was not statistically significant ($p=0.116$ and $p=0.239$, respectively). The details of the preliminary results for student learning are reported elsewhere (Granger et al., 2009).

In general, the results indicate that the effects of professional development plus teaching with the GEMS *SSCS* were positive for teachers as compared with the control group. The estimated GEMS main effects for the teacher outcomes are summarized in Table 2. The results for each outcome were obtained using a regression model containing the coded *GEMS* treatment variable and the corresponding pre-measure of the outcome. For example, the estimated model for teacher self efficacy (Efficacy) outcome (*EF*) was

$$EF = 96.6 + 8.01\ GEMS + 0.200\ \mathrm{Pr}e\_EF \qquad (1)$$

where *Pre_EF* is the teacher pre-Efficacy variable. From (1), it is seen that the predicted outcome for the GEMS treatment group was 8.01 units higher than that for the control group, adjusting for any group differences with respect to *Pre_EF*. Results associated with this model, given in the first row of Table 2, include the unstandardized GEMS effect estimate, the standard error of this estimate, the p-value of the effect, and the standardized GEMS effect. The latter was obtained by dividing the unstandardized effect by the standard deviation of the outcome.

Table 2. GEMS Main Effects for Teacher Outcomes

Outcome	Unstandardized GEMS Effect	Standard Error	p-value	Standardized GEMS Effect[a]
Self Efficacy	8.01	6.50	0.223	0.33
Outcome Expectancy	8.11	6.19	0.196	0.35
Achievement	3.10*	0.579	< 0.001	0.86*
Confidence	0.037*	0.177	0.089	0.38*
Attitude	-5.50	16.0	0.733	-0.11

* Statistically significant at the 0.010 level.[a] Standardized effects have been obtained by dividing the unstandardized GEMS coefficient by the outcome standard deviation. The standard deviations for the above outcomes are, in order: 24.5, 23.3, 3.59, 0.809, 6.43, 5.38, and 50.0

Considering all of the outcomes in Table 2, the GEMS effects were positive and statistically significant for the Achievement (Space Science Content test) and Confidence (items about teacher confidence in their content knowledge were included at the end of the Space Science Content test). Using Cohen's (1977) characterizations of the strength of standardized effects, the effects ranged from approximately medium to large in strength. There was no support for GEMS main effects for the Attitude outcome. The GEMS effects for self efficacy and outcome expectancy were positive but not statistically significant at the 0.10 level. Note, however, that in the discussion below of effect moderation, it will be concluded that there were GEMS effects for self efficacy and outcome expectancy for particular subgroups of teachers.

The search for possible teacher moderator variables (see Introduction for discussion of teacher moderator variables) involved the extension of the main effect models (like that in [1] above) to include a product term representing the interaction between *GEMS* and the pre-measure. Consider, for example, the following estimated interactive model for the self efficacy outcome

$$EF = -0.1 + 145.8\ GEMS + 0.949\ Pre\_EF - 1.075\ GEMS * Pre\_EF \qquad (2)$$

As indicated in the first row of Table 3, the interaction coefficient of -1.07 was statistically significant (p = 0.016) and indicated that the GEMS treatment versus control contrast decreased by 1.07 units when the *Pre_EF* variable increased by one unit.

Table 3: Moderators of the GEMS Effect on Outcomes

Outcome	Moderator	Simple Effect Expression	Standardized Simple Effects[b]

	(with label and p-value)	(Standardized form in parentheses)[a]	Low on Moderator	Average on Moderator	High on Moderator
Self Efficacy	Pre-Efficacy (*Pre_EF*) (p = 0.016)	145.8 – 1.075**Pre_EF* (5.95 – 0.0439**Pre_EF*)	1.04	0.35	-0.34
	Pre-Attitude (*Pre_AT*) (p = 0.035)	121.4 – 0.432**Pre_AT* (4.96 – 0.0176**Pre_AT*)	1.35	0.49	-0.37
Outcome Expectancy	Pre-Outcome Expectancy (*Pre_OE*) (p =0.069)	106.0 – 0.762**Pre-OE* (4.55 – 0.0327**Pre_OE*)	0.88	0.36	-0.16

[a] The outcome standard deviations used for the standardization for the Self Efficacy and Outcome Expectancy outcomes were 24.5 and 23.3, respectively.

[c] The simple effects have been computed for: teachers one standard deviation below the moderator mean (Low on Moderator), teachers at the moderator mean (Average on Moderator), and teachers one standard deviation above the moderator mean (High on Moderator). The means (and standard deviations) of the Pre_Efficacy, Pre_Attitude, and Pre_Outcome Expectancy variables were: 127.6 (15.8), 253.4 (48.7), and 128.2 (15.9), respectively.

The model in (2) can be decomposed into two models, one for each of the two study groups, by substituting the coding values of 1 and 0 into the *GEMS* variable to result in:

$$\begin{aligned} &GEMS\ Group\ (GEM = 1): EF = 145.7 - 0.126\ \mathrm{Pr}e\_EF \\ &Control\ Group\ (GEMS = 0): EF = -0.1 + 0.949\ \mathrm{Pr}e\_EF \end{aligned} \quad (3)$$

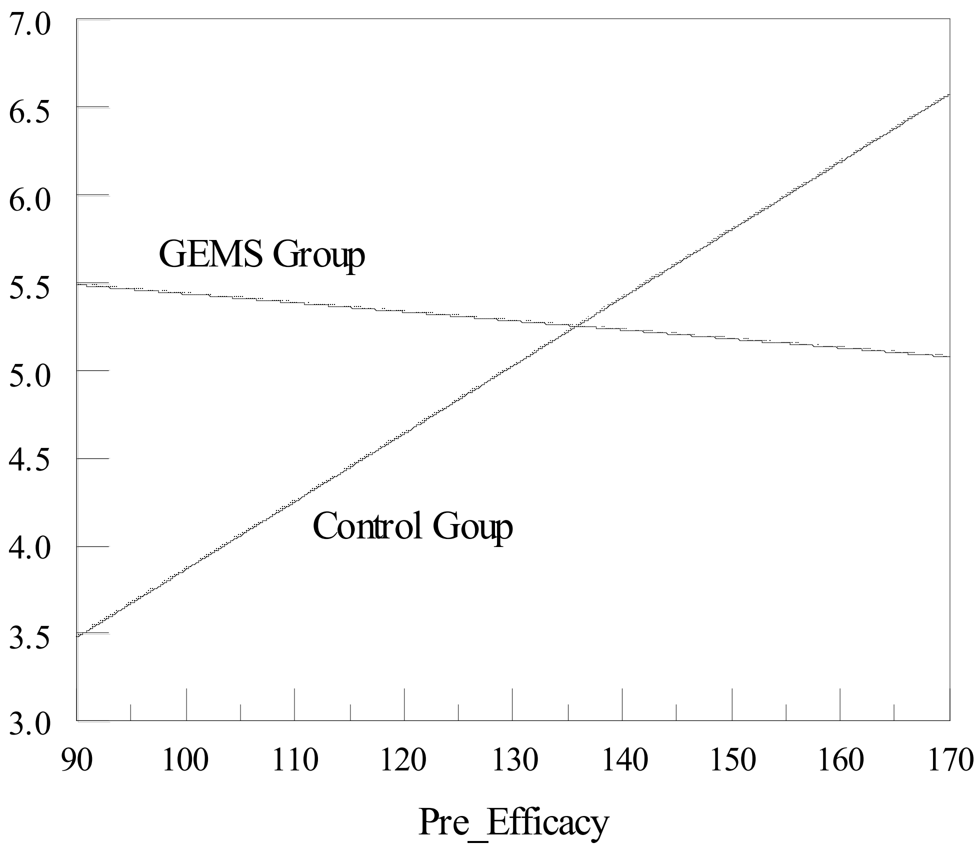

Figure 1: Predicted Self efficacy outcome plotted as a function of the moderator variable for the experimental and control groups. The GEMS standardized simple effect (i.e., the standardized difference between the predicted outcomes for the GEMS and control groups) is the vertical separation of the two lines.

Figure 1 shows the predicted self efficacy outcome plotted as a function of the Pre_Efficacy moderator for each of the two groups over the approximate range of the moderator. Each model has been standardized for the figure, dividing each coefficient by the standard deviation of the outcome. The standardized simple effect of GEMS is represented in Figure 1 by the vertical separation of the two predicted value lines for given values of Pre_Efficacy. It is seen that for very low values of the moderator, the standardized GEMS effect is positive and quite large. As the moderator increases, the size of the positive effect decreases until, at a moderator value of approximately 135, the estimated simple effect is zero. As the moderator continues to increase, the estimated standardized GEMS simple effect becomes negative.

An expression for the simple effect is obtained by identifying all terms in (2) involving *GEMS* and dividing out the *GEMS* variable (or, mathematically, by finding the partial derivative of [2] with respect to *GEMS*), resulting in

$$\textit{Simple effect of } GEMS = 145.8 - 1.075 \ \mathrm{Pr}e\_EF \quad (4)$$

After standardizing, this estimated simple effect is represented graphically as the straight dashed line in Figure 2. As noted in the discussion of Figure 1, the simple effect was positive and large for low values of the Pre_Efficacy moderator, decreased in magnitude with increasing values of the moderator, and became negative for higher moderator values. A 90% confidence band on the estimated effect is shown in Figure 2 by the two curved lines (e.g., Tate, 2004). For each value of Pre_Efficacy, the associated estimated GEMS effect was statistically significant if the corresponding two values in the confidence band did not capture zero. Thus, the estimated effects were statistically significant for Pre_Efficacy values up to approximately 121; for higher values of the moderator all effects were statistically insignificant.

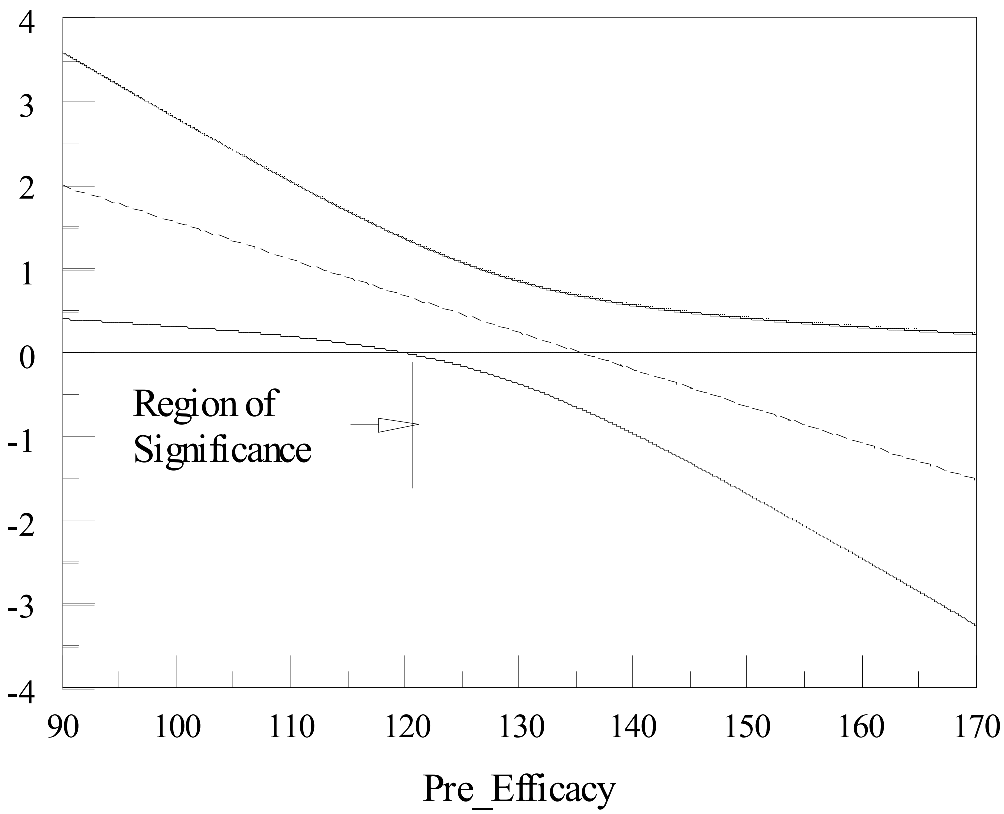

Figure 2: Estimate of the standardized simple effect of GEMS on the Self efficacy outcome as a function of pre-Efficacy (straight dashed line). The two solid curves represent the 90% confidence band on the effect estimate. The estimated effects are statistically significant at the 0.10 level for all values of Pre-Efficacy where the confidence interval does not capture zero.

The first row of Table 3 summarizes this information for Pre_Efficacy as a moderator of GEMS for the self efficacy outcome. Both the unstandardized and standardized forms of the simple effect expression are given, and the variation of the simple effect over the range of the moderator is represented by the three standardized effects shown for low, average, and high values of Pre_Efficacy. The remainder of Table 3 summarizes results for other moderators that were identified. For example, a preplanned interaction hypothesis stating that Pre-Attitude was a moderator for the

Efficacy outcome was tested and supported. The second row of Table 2 shows an interaction that behaved very much like that discussed above for the Pre-Efficacy moderator. The simple effect of GEMS was positive and large for low values of Pre-Attitude and negative for high Pre-Attitude values. (Given the support for two different moderators for the self efficacy outcome, Pre_Efficacy and Pre_Attitude, another analysis added both as moderators into the same model. Dropping nonsignificant terms, the resulting model contained only the Pre_Efficacy moderator, not the Pre_Attitude moderator.) Finally, the Pre_Outcome Expectancy moderator for the Outcome Expectancy outcome was found to be statistically significant, exhibiting an interactive pattern very similar to that associated with that for the self efficacy outcome. Given evidence that self efficacy and Outcome Expectancy were measuring closely related constructs, this result was not surprising. There was no evidence for any other moderators for the seven teacher outcomes. Confidence interval diagrams for the second and third rows of Table 3 closely resemble that of the first row (shown in Figure 2) and have not been included herein.

In summary, teacher achievement on the post content test for the GEMS group compared to the control group was positive and statistically significant ($p<.001$). Teacher confidence with their content knowledge on the post assessment for the GEMS group compared to the control group was positive and statistically significant ($p=.089$). When the self efficacy scores were examined in light of pre-test self efficacy scores as a moderator variable for the GEMS effects, there was a positive and statistically significant effect on post self-efficacy for the GEMS group compared to the control group for teachers who had a low pre-test self efficacy score, a positive but smaller non-significant effect for those that had a moderate pre-test self efficacy score, and a small negative non-significant effect for those that had a high pre-test self efficacy score. When the self efficacy scores were examined in light of pre-test attitudes about science scores as a moderator variable for the GEMS effects, there was a positive and statistically significant effect on post self efficacy for the GEMS group compared to the control group for teachers who had a low pre-test attitude score, a positive but smaller non-significant effect for those that had a moderate pre-test attitude score, and a small negative non-significant effect for those that had a high pre-test attitude score.

When the teaching outcome expectancy scores were examined in light of pre-test teaching outcome expectancy scores as a moderator variable for the GEMS effects, there was a positive and statistically significant effect on the post outcome expectancy score for the GEMS group compared to the control group for teachers who had a low pre-test teaching outcome expectancy score, a positive but smaller non-significant effect for those that had a moderate pre-test teaching outcome expectancy score, and a small negative non-significant effect for those that had a high pre-test teaching outcome expectancy score.

Thus, the analyses for main effects and moderators supported the presence of positive GEMS effects for all teacher outcomes except Attitudes toward science, though attitude was a moderator for teacher with low pre-self efficacy scores. When interactions were supported, the GEMS effects were largest for teachers at lower levels of the moderators.

Discussion

It is important to note at the outset of this discussion that this is a preliminary analysis of the results from the first year of a two-year study. The completed study will consist of two cohorts of teachers (year one and year two) who participated in the same regimen of professional development and teaching. The experimental methods were the same for each cohort.

Although more research into the preparation of new teachers is critical and necessary, research into how to best allow for continued professional development of practicing teachers is vitally important if we are to move toward the vision of teaching called for in science education reforms. Equally important is research into ways to support teacher in the implementation of what they learn during professional development when they return to the classroom. Central to this implementation is the identification of appropriate curricula that support teachers, for even with high-quality post professional development follow-up and science coaching, ultimately teaching practice is carried out by individual teachers and their practice is shaped through interaction with the curriculum on a day-to-day basis.

From teacher professional development studies we know that some of the most significant influences on student learning are the skills and abilities of the teachers (e.g., content knowledge and pedagogical content knowledge) (Blanchard et al., in review; Cohen & Hull, 1998; Goe, 2007; Schneider, 2005). Too, research points to teacher beliefs about how students learn, students' abilities and capabilities, the teacher's role in instruction, and their own self efficacy, self doubt, and attitude (among others) as central to the adoption of new instructional practices (e.g., Ashton & Webb, 1986; Bandura, 1977; Cronin-Jones, 1991; Krajcik, et al., 1994; Rice & Roychoudhury, 2003; Ross, 1992; Settlage et al., 2008; Tschannen-Moran et al., 1998; Wheatley, 2002; Woolfolk et al., 1990). Of particular relevance to this study, teacher self efficacy centers around a teacher's beliefs about their own ability to influence student learning and their ability to enact specific instructional interventions (e.g., Bandura, 1977; Settlage, et al., 2009). Outcome expectancy in the context of teaching reflects a more general view of the possible effects of teaching on learning; for example, high outcome expectancy reflects a belief that effective instruction can overcome other factors that negatively influence student learning (Bandura, 1977, 1997; Settlage et al., 2009). It has been further suggested that at least some amount of self doubt with respect to teaching and curriculum enactment is not necessarily a negative force as it may stimulate professional growth and development (Settlage et al., 2009; Wheatley, 2002; Woodbury & Gess-Newsome, 2002).

The findings reported herein suggest that well-designed, reform-based curriculum, such as the GEMS *SSCS*, combined with professional development centered around that curriculum, can support the development of elementary teachers' content knowledge. As discussed above, teacher content knowledge has been shown to be a necessary but not sufficient factor influencing science instruction. Our results further indicate that use of the GEMS curriculum in combination with professional development supported teachers' confidence with their content knowledge. This increased confidence and content knowledge may also exert some influence on science teaching, though our work does not delve more deeply into this idea. As many practicing elementary teachers report lack of content knowledge as necessary to the effective teaching science, especially in a reformed manner (e.g., authors E.G. and T.B., personal experience during > 12 combined years of professional development work, Lee et al., 2003; Luera et al., 2005), improving content knowledge and teachers' confidence in their knowledge are features that should be a part of any professional development effort.

Our findings further suggest that well-designed, reform-based curricula in combination with professional development can be effective in shaping teachers' beliefs about teaching. The evidence provided herein supports the recursive relationship between context and knowledge and beliefs described by others (e.g., Roehrig & Kruse, 2005; Smith & Southerland, 2007; Yore et al., 2007). Further, this study suggests that the GEMS *SSCS* curriculum materials and professional development were more effective in supporting teachers who had more to learn, that is, those with lower self efficacy and outcome expectance at the outset of the study. This effect lessened for teachers who had high self efficacy and outcome expectancy at the outset.

Several features of the GEMS curriculum and professional development are potential contributors to the modification of teacher beliefs as a result of this experience. The GEMS *SSCS* is structured to provide experiences that confront common space science misconceptions including frequent opportunities to talk about these experiences including opportunities to provide evidence for the correct conceptions. Through the professional development the teachers first experienced these activities and discussions as learners, just as their students would, then were provided with further experiences to help them delve more deeply into the subject matter, and finally experiences to help them understand the pedagogy employed by the curriculum. Following this they taught the unit which provided them with further experience with the content. Finally, there is a very well-written content background for the teachers section included in the curriculum. Thus, teachers who initially had little space science content knowledge became much more knowledgeable and secure with what they knew about the topics in the unit. It may also be that, for those still somewhat unsure of themselves, the scripted nature of the curriculum functions as another support.

Our findings further call into question the role of initial high self efficacy in hindering teacher development echoing suggestions in the literature (e.g., Settlage et al., 2009; Wheatley, 2002). That is, a measure of self doubt has been postulated by these authors and others as a mechanism for motivating change in teaching practice. As cited by Settlage, et al (2009), according to Dewey (1916) such self doubt "…marks an inquiring, hunting, searching attitude, instead of one of mastery and possession. Through its critical process true knowledge is revised and extended, and our convictions as to the state of things reorganized." However, this study does not specifically explore the role of self doubt. That is, we cannot determine from our data whether a lack of self doubt contributed to our findings for teachers with high initial self efficacy and outcome expectancy, whether these results reflected a ceiling effect on the self efficacy and outcome expectancy variables, or whether some other factor(s) influenced these results.

Finally, teachers' attitudes about science at the outset of the study also moderated their final self efficacy. That is, those that initially had poor attitudes about science showed significant improvement in their self efficacy following professional development and teaching with the GEMS curriculum. This provides further evidence supporting the positive effects of the GEMS *SSCS* curriculum and professional development on teacher beliefs about teaching.

Implications

Our results indicate that well-designed curriculum and professional development around it *can* be effective in shaping teachers professional development in science and the teaching practices associated with reformed science education instruction. Further, our results suggest that what teachers learn from professional development and teaching with the curriculum varies according to what their knowledge, beliefs, and/or attitudes are entering the experience. Those working in professional development need to be very mindful of the needs and abilities of the teachers that enter such experiences. Many district personnel, not familiar with teaching as an act of interpreting materials, may forgo the need for professional development with new curricular materials that is centered around both the materials *and* the pedagogy that underpins them.

Future Research

The research reported herein examined the combination of professional development and well-designed, reform-based curricula on teachers' knowledge and beliefs. Further research is needed to examine the effectiveness of implementing the curriculum without professional development on teachers' knowledge and beliefs in order to shape the practices of school administrations.

Our results do not allow us to tease apart which features of the GEMS *SSCS* curriculum and/or professional development were the most salient in supporting the development of teachers and this is an avenue for future research. Finally, whether such curriculum materials are as effective in supporting the professional development of teachers in the secondary grades (who often are perceived as having greater content knowledge and/or sophistication with science teaching practices) is another area for future research.

Acknowledgements: We are deeply indebted to Dr. Richard Tate, our statistician, for his talent, help, instruction, and advice with the statistical analyses in this paper and to our research assistants, Francesca Gerbino and Katrina Dial for their assistance with data management. The staff of the Florida Center for Research in STEM Education was essential in helping us to administer this project. We wish to acknowledge the significant contribution of our science coach, Tracie Ragland, to the project. Finally, the project is indebted to the teachers who agreed to constrain their teaching to the assigned methodology and curriculum in order to further knowledge about their profession.

This research was supported by a grant from the Florida Center for Research in Science, Technology, Engineering and Mathematics education, (FL DOE # 371 – 96700- 7SF01, 371 -96700 – 8SS01, 371-96700-9SF01).

References

Abrams, E., Southerland, S.A., & Evans, C. (2007). Inquiry in the classroom: The necessary components of a useful definition. In E. Abrams, S.A. Southerland, & P Silva (Eds.), *Inquiry in the classroom: Realities and Opportunities*. Greenwich, CT: Information Age Publishing.

American Association for the Advancement of Science. (1993). *Benchmarks for Science Literacy*, New York: Oxford University Press.

Ashton, P.T., & Webb, R.B. (1986). *Making a Difference: Teachers' Sense of Self Efficacy and Student Achievement.* New York: Longman.

Bandura, A. (1977). Self Efficacy: Toward a Unifying Theory of Behavioral Change. *Psychological Review*, 84, 191-215.

Bandura, A. (1997). *Self Efficacy: The Exercise of Control*. New York: W.H. Freeman.

Barnett, M., & Morran, J. (2002). Addressing Children's Alternative Frameworks of the Moon's Phases and Eclipses. *International Journal of Science Education*. *24*, 859-879.

Blanchard, M., Southerland, S.A., & Granger, D.E. (2009). No silver bullet for inquiry: Making sense of teacher change following an inquiry-based research experience for teachers. *Science Education, 93*(2), 322-360.

Blanchard, M., Southerland, S.A., Sampson, V., Annetta, L., & Granger, D.E. (In Review). Student learning in a laboratory setting: Comparing the effectiveness of inquiry-based versus deductive science teaching methods. *Science Education.*

Bybee, R. W. (1997). *Achieving science literacy: From Purposes to Practices*. Portsmouth, NH. Heinemann.

Chinn, C., & Maholtra, B. (2002). Epistemologically Authentic Inquiry in Schools: A Theoretical Framework for Evaluating Inquiry Tasks. *Science Education, 86*, 175-218.

Cohen, J. (1977). *Statistical Power Analysis for the Behavioral Sciences (Revised Edition)*. New York: Academic Press.

Cohen, D.K., & Hill, H.C. (1998). *Instructional policy and classroom performance: The mathematics reform in California (CPRE Research Report No. RR-39). Philadelphia: Consortium for Policy Research in Education.*

Collaboration for Astronomy Education Research (CAER). (1999). Astronomy Diagnostic Test. http://solar.physics.montana.edu/aae/adt.

Colburn, A. (2000). An inquiry primer. *Science Scope*, *23 (6),* 42-44.

Crawford, B.A. (2000). Embracing the Essence of Inquiry: New Roles for Science Teachers. *Journal of Research in Science Teaching, 37*, 916-937.

Crawford, B.A. (2007). Learning to Teach Science in the Rough and Tumble of Practice. *Journal of Research in Science Teaching, 44,* 613-642.

Cronin-Jones, L.L. (1991). Science Teacher Beliefs and Their Influence on Curriculum Implementation: Two Case Studies. *Journal of Research in Science Teaching, 28*, 235-250.

Dewey, J. (1916). *Democracy and Education*. New York: McMillan.

Dira-Smolleck, L.A. (2004). *The development and validation of an instrument to measure preservice teachers' self-efficacy in regards to the teaching of science as inquiry*. Unpublished doctoral dissertation. The Pennsylvania State University. (Creative Commons Attribution 2.5 License).

Duschl, R., Schwingruber, H.A., & Shouse, A.W. (2007) *Taking Science to School*: *Learning and Teaching Science in Grades K-8*. Committee on Science Learning, Washington, D.C.: National Academy Press.

Elmore, R. F. & D. Burney (1997). *Investing in teacher learning: Staff development and instructional improvement in community district #2*, New York City, National Commission on Teaching and America's Future and Consortium for Policy Research in Education.

Feldman, A., Divoll, K., & Rogan-Klyve, A. (2009). Research education of new scientists: Implications for science teacher education. *Journal of Research in Science Teaching, 46(4), 442-459.*

Flick, F. B. (2003, March). *Inquiry as cognitive process*. Paper presented at the annual meeting of the National Association for Research in Science Teaching, Philadelphia.

Fraser, B.J. (1978). *Test of Science Related Attitudes*. Australian Council for Educational Research. Camberwell, Victoria, Australia: Allenby Press.

Gess-Newsome, J., Southerland, S.A., Johnston, A. & Woodbury, S. (2003). Educational reform, personal practical theories, and dissatisfaction: The Anatomy of change in college science teaching. *American Educational Research Journal, 40*(3), 731-767.

Glenn Commission. (2000). *Before It's Too Late*, a report to the nation from the National Commission on Mathematics and Science Teaching for the 21st Century. Jessup, MD: U.S. Department of Education, Education Publication Center.

Goe, L. (2007). *The link between teacher quality and student outcomes: A research synthesis.* National Center for Teacher Quality. Nashville, TN:Vanderbuilt University. Retrieved April 7, 2009 from http://www.tqsource.org/link.php.

Granger, E.M., Bevis, T.H., Saka, Y., & Southerland, S. (2009). Comparing the Efficacy of Reform Based and Traditional/Verification Curricula to Support Student Learning about Space Science. Paper presented at the annual meeting of the National Association for Research in Science Teaching, Garden Grove, CA.

Howard, M. & Associates. (2007). MSP Teacher Feedback on Science Instructional Materials. (Instrument developed based on others used in National Science Foundation projects).

Karplus, R., & Their, H. D. (1967). *A new look at elementary school scienc*e. Chicago: Rand McNally.

Kavanagh, C., & Sneider, C. (2007a) Learning about gravity I. Free Fall: A guide for teachers and curriculum developers. *Astronomy Education Review, 5* (2).

Kavanagh, C. & Sneider, C. (2007b) Learning about gravity II. Trajectories and orbits: A guide for teachers and curriculum developers, *Astronomy Education Review, 5* (2).

Klahr, D., & Nigam, M. (2004). The equivalence of learning paths in early science instruction: Effects of direct instruction and discovery learning. *Psychological Science, 15,* 661–667.

Krajcik, J.S., Blumenfeld, P.C., Marx, S.W., & Soloway, E. (1994). A Collaborative Model for Helping Middle Grade Science Teachers Learn Project Based Instruction. *The Elementary School Journal, 94,* 483-497.

LaChance, A.M. (2004). Using Standards Based Mathematics Curricula to Better Prepare Mathematics Teachers. Paper presented at the North American Chapter of the International Group for the Psychology of Mathematics Education, Delta Chelsea Hotel, Toronto, Ontario Canada. PDF 2001-02-05 from http://www.allacademic.com/meta/p117665_index.html.

Lawrence Hall of Science. (2007). *Space Science Curriculum Sequence.* Great Explorations in Math and Science. Berkeley, CA: University of California-Berkeley Press.

Lederman, J. & Lederman, N. (2005*). Developing and Assessing Elementary Teachers' and Students' Understandings of Nature of Science and Scientific Inquiry.* Paper Presented at the Annual Meeting of the National Association for Research in Science Teaching, Dallas, Texas, April, 2005.

Lederman, N., Lederman, J., Wickman, P. - O. & Lager-Nyqvist, L. (2007, April). *Development of a Valid and Reliable Protocol for the Assessment of Early Childhood Students' Conceptions of Nature of Science and Scientific Inquiry.* Paper presented at the annual conference of the National Association for Research in Science Teaching, New Orleans, LA.

Lee, C. & Houseal, A. (2003). Self-efficacy, Standards, and Benchmarks as Factors in Teaching Elementary School Science. *Journal of Elementary Science Education*.

Leura, G., Moyer, R.H., & Everett, S.A. (2005) What Type and Level of Science Content Knowledge of Elementary Education Students Affects their Ability to Construct an Inquiry-Based Science Lesson? *Journal of Elementary Science Education*.

Luft, J. (2007). Minding the Gap: Needed Research in Beginning/Newly Qualified Science Teachers. *Journal of Research in Science Teaching, 44 (4),* 532-537.

Luft, J.A., Roehrig, G.H., Brooks, T., & Austin, B.A. (2003). *Exploring the Beliefs of Secondary Science Teachers Through the Interview Maps*. Paper presented at the annual meeting of the National Association for Research in Science Teaching, Philadelphia, PA.

Marx, R. W., Blumenfeld, P. C., Krajcik, J. S., Fishman, B., Soloway, E., Geier, R., & Tal, R. T. (2004). Inquiry-based science in the middle grades: Assessment of learning in urban systemic reform. *Journal of Research in Science Teaching, 41*, 1063-1080.

McIntyre, D. & Hagger, H. (1992) Professional development through the Oxford Internship Model, *British Journal of Educational Studies*, *40*, 264-283.

Metz, K. E. (2004). Children's understanding of scientific inquiry: Their conceptualization of uncertainty in investigations. *Cognition and Instruction*, *22*, 219-290.

National Academy of Sciences, Committee on Prospering in the Global Economy of the 21st Century. (2007). *Rising Above the Gathering Storm*, Washington, D.C. National Academies Press.

National Research Council. (1996). *National Science Education Standards*. Washington, D.C., National Academy Press.

National Research Council. (2000). *Inquiry and the National Science Education Standards*. Washington, D.C.: National Academy Press.

National Research Council. (2008). *Ready, Set, Science: Putting Research to Work in K-8 Science Classrooms.* Washington, D.C.: National Academy Press.

National Science Board, Committee on Education and Human Resources. (2006). America's Pressing Challenge—Building a Stronger Foundation. Arlington, VA, National Science Foundation.

Pringle, R. M., & Carrier, S. M. (2005). The potential impacts of upcoming high-stakes testing on the teaching of science in elementary classrooms. *Research in Science Education, 35,* 347-361.

Rice, D.C., & Roychoudhury, A. (2003). Preparing More Confident Preservice Elementary Science Teachers: One Elementary Science Methods Teacher's Self-Study. *Journal of Science Teacher Education, 14,* 97-126.

Ross, J.A. (1992). Teacher Efficacy and the Effect of Coaching on Student Achievement. *Canadian Journal of Education, 17,* 51-65.

Sawada, D., Piburn, M.D., Kyle, W.C., & Alport, J.M. (2002) Measuring Reform Practices in Science and Mathematics Classrooms: The Reformed Teaching Observation Protocol, *School Science and Mathematics, 102*, 245-253.

Sadler, T. D. (2006). Promoting discourse and argumentation in science teacher education. *Journal of Science Teacher Education*, 17(4), 323-346.

Sadler, P., Coyle, H., Cook-Smith, N. & Miller, J. (2007). Misconceptions-oriented Standards-based Assessment Resources for Teachers (MOSART), Cambridge, MA: Harvard College. http://www.cfa.harvard.edu/smgphp/mosart/site_map.html.

Sampson, V., & Benton, A. (2006). Development and validation of the Beliefs about Science Teaching and Learning (BARSTL) Questionnaire. Paper presented at the annual meeting of the Association for Science Teacher Education, Portland, OR.

Schneider, R.M., Krajcik, J., & Blumenfeld, P. (2005). Enacting Reform-based Science Materials, *Journal of Research in Science Teaching, 42 (3)*, 283-312.

Settlage, J., & Blanchard, M. (2007). Requisite Teacher Knowledge about Inquiry: Resources for Thinking about Teaching for Inquiry. In E. Abrams, S.A. Southerland, & P Silva (Eds.), *Inquiry in the classroom: Realities and Opportunities*. Greenwich, CT: Information Age Publishing.

Settlage, J., & Meadows, L. (2002). Standards-based reform and its unintended consequences: Implications for science education within America's urban schools. *Journal of Research in Science Teaching*, 39(2), 114-127.

Settlage, J., Southerland, S.A., Smith, L., & Ceglie, R. (2009). Constructing a doubt-free teaching self: Self-efficacy, teacher identity, and science instruction within diverse settings. *Journal of Research in Science Teaching, 46*(1), 102-125.

Smith, L.K. (2005). The impact of early life history on teachers' beliefs: In-school and out-of-school experiences as learners and knowers of science. *Teachers and Teaching: Theory and Practice, 11*(1), 5-36.

Smith, L., & Southerland, S.A. (2007). Reforming practice or modifying reforms? Elementary teachers' response to the tools of reform. *Journal of Research in Science Teaching, 43*(3), 396-423.

Southerland, S.A., Smith, L. K., Sowell, S., & Kittleson, J. (2007). Resisting unlearning: Understanding science education's response to the United State's national accountability movement. *Review of Research in Education, 31*, 45-77.

Tate, R. (2004). Interpreting Hierarchical Linear and Hierarchical Generalized Linear Models with Slopes as Outcomes. *Journal of Experimental Education, 73*, 71-95.

Tschannen-Moran, M., Hoy, A.W., & Hoy, W.K. (1998). Teacher Efficacy: Its Meaning and Measure. *Review of Educational Research*. 68, 202-248.

Warrington, M., Younger, M., Williams, J. (2000) Student Attitude, Image and the Gender Gap, *British Educational Research Journal*, 26 (1), 393-407.

Wheatley, K.F. (2002) The Potential Benefits of Teacher Efficacy Doubts for Educational Reform, *Teaching and Teacher Education, 18*, 5-22.

Windschitl, M. (2004) Folk Theories of "Inquiry:" How Preservice Teachers Reproduce the Discourse and Practices of an Atheoretical Scientific Model. *Journal of Research in Science Teaching, 41*, 481-512.

Woodbury, S., &Gess-Newsome, J. (2002) Overcoming the Paradox of Change Without Difference: A Model of Change in the Arena of Fundamental School Reform, *Educational Policy, 16*, 763-782.

Woolfolk, A.E., & Hoy, W.K. (1990) Prospective Teachers' Sense of Efficacy and Beliefs About Control. *Journal of Educational Psychology, 82*, 81-91.

Yager, R. E. (1992). Viewpoint: What we did not learn from the 60's about science curriculum reform. *Journal of Research in Science Teaching*, *29*(8), 905-910.

Yore, L, Henriques, L., Crawford, B., Smith, L.K., Zwiep, S., & Tillotson, J. (2008). Selecting and using inquiry approaches to teach science: The influence of context in elementary, middle & secondary schools. In E. Abrams, S.A. Southerland, & P. Silva (Eds.). Inquiry in the classrooms: Challenges and Opportunities. Greenwich, Ct: Information Age Publishing.