

## Essential Aspects of Science Teacher Professional Development: Making Research Participation Instructionally Effective

**Sherry A. Southerland**  
**Ellen M. Granger**

*Florida State University*

**Roxanne Hughes**

*National High Magnetic Field Laboratory*

**Patrick Enderle**

*Georgia State University*

**Fengfeng Ke**

*Florida State University*

**Katrina Roseler**

*Chaminade University of Honolulu*

**Yavuz Saka**

*Bulent Ecevit University*

**Miray Tekkumru-Kisa**

*Florida State University*

*Current reform efforts in science place a premium on student sense making and participation in the practices of science. Given the disparity between these activities and current teaching practices, effective means of professional development around such practices must be identified. We use a close examination of 106 science teachers participating in Research Experiences for Teachers (RET) to identify, through structural equation modeling, the essential features in supporting teacher learning from these experiences. Findings suggest that participation in RET shape science teacher practice and beliefs, which in turn influence practice. Essential features of RET include engaging teachers socially in the research context and in research projects that are personally relevant to them. The model suggests ways to maximize the professional development potential of RET intended to support engagement in disciplinary practices.*

**Keywords:** *professional development, science education, structural equation modeling, teacher education/development*

A quiet revolution is occurring in the learning goals that scientists and science educators have set for students. Scientific literacy, an ambiguously defined construct, has given way to the goal of students becoming proficient in science, which involves more than an understanding of important concepts; it centers on being able to do science. From this vantage point, doing science focuses on students engaging in productive sense making about the natural world (National Research Council [NRC], 2014). With that goal, the Next Generation Science Standards (NGSS) are performance expectations that integrate three dimensions of science learning: core ideas, scientific practices, and cross-cutting concepts (Krajcik, Codere, Dahsah, Bayer, & Mun,

2014). These performance expectations are meant to reflect the disciplinary practices of science to engage students in productive sense making as a vehicle to support science learning (National Academies of Sciences, Engineering, & Medicine [NAS] 2015). Indeed, a central component of the NGSS in terms of vision of science learning is that students' science learning is intimately tied to their engagement in investigations involving phenomena.

If learning science is to result from productive sense making, there must be a “fundamental change in the way science is taught” (NAS, 2015, p. 1) and envisioned by teachers. If teachers are to support students in framing their role in science classrooms as one of sense making instead



of assignment completion (Passmore, 2014), they will need to develop a “deep craft knowledge” for how to approach science as a knowledge-building activity (Moon, Michaels, & Reiser, 2012). It has been widely argued that teachers themselves need to actively engage in the practices of science that lead to productive sense making so that they may develop a fluency that can come only from “continuing and extensive research experiences” (NAS, 2015, p. 100).

Few science teachers have had such research experiences, and much of the undergraduate preparation for science teachers precludes authentic research experiences (Crawford, 2000; Windschitl, 2003). Thus, if the goals of the NGSS are to be realized, professional development (PD) that involves teachers in scientific research is needed. As defined by Desimone (2011), an emerging and broader view of teacher PD considers teacher learning as being interactive, social, and based in a community of practice. In fact, Little (1993) argued that the dominant training model of PD, which mainly focuses on expanding individual repertoire of skillful classroom practice, is not enough. One of the alternative models that she discussed is participation in special institutes, such as summer institutes sponsored by the National Science Foundation (NSF). As stated by Little, “judging by teachers’ accounts, such institutes and centers offer substantive depth and focus, adequate time to grapple with ideas and materials, the sense of doing real work rather than being ‘talked at’” (p. 137). NSF-funded Research Experiences for Teachers (RET) programs, which can be considered an example of such institutes, provide STEM teachers (i.e., science, technology, engineering, and mathematics) with opportunities to collaborate with university faculty, and they encourage them to translate research into classroom (NSF, 2016). PD “can be viewed as the ongoing learning experience of a teacher. . . . For science teachers, the specific character of science is an essential ingredient of teacher practice, of teacher learning, and thus of programs designed to facilitate these outcomes” (Luft & Hewson, 2014, p. 889). Science teachers have access to a “carnival of options” for PD (Wilson, 2013, p. 311); however, creating and/or selecting the most effective form of PD represents a huge undertaking. Scant attention has been paid to understanding the relative effectiveness of such programs or the mechanisms through which these programs exert their influence (Banilower et al., 2013), beyond teacher self-report (Capps, Crawford, & Constat, 2012).

Wilson (2013) and others (e.g., NAS, 2015) have argued that if the goals of the NGSS are to be realized, the field must create, study, and refine models of PD that can be successful in supporting teachers’ disciplinary engagement in science so that this engagement can inform their teaching. This research examines two examples of one long-standing form of PD designed to engage teachers in the practices of science, RET, in an attempt to identify the ways in which teacher learning is fostered through research experiences

with scientists and to highlight the essential aspects of these experiences that facilitate learning.

## Related Literature

### *Teacher Thinking and PD*

The teacher-centered systemic reform framework (Gess-Newsome, Southerland, Johnston & Woodbury, 2003), represented in Figure 1 provided the theoretical grounding for this study by acknowledging the complex nature of multiple influences on teachers’ instructional practice. This framework posits that teacher thinking and affect are central to any systemic reform effort, as they are critical filters that teachers use to interpret the many messages they receive through personal experiences and contextual forces. Furthermore, teachers’ thinking and affect reciprocally interact with teacher practice, each shaping the other and both serving as critical targets for engendering meaningful change in the classroom. What teachers know and what teachers believe greatly influence how they teach and how they “take up” new messages regarding their teaching, a process that is heavily influenced by a teacher’s thinking as one enters the PD experience (Gregoire, 2003; Jones & Carter, 2007; Roehrig & Kruse, 2005). PD can be understood to exert its influence by shaping teachers’ thinking, which in turn shapes their practice.

From the wide range of features of teacher thinking examined in the past research, three constructs emerge as fundamentally important to examine: teaching self-efficacy, pedagogical discontentment, and teacher beliefs about science teaching and learning. Teaching self-efficacy describes a teacher’s sense of his or her abilities to teach science in general. “Perceived self-efficacy is defined as people’s judgments of operative capabilities to organize and execute courses of action required to attain designated types of performances” (Bandura, 1986, p. 391). The work of Pajares (1997), Bandura (1986), and others (Granger, Bevis, Saka, Southerland, & Sampson, 2012; Gregoire, 2003; Wheatley, 2002) have identified that a teacher’s sense of self-efficacy is an influential mediator for change within the context of reform, as a moderate sense of self-efficacy has been identified as essential for a teacher’s “uptake” of new instructional practice. These and other studies also demonstrate that self-efficacy as it relates to inquiry is important to changes in teaching practice, specifically on teachers’ abilities to teach through inquiry (e.g., Smolleck, Zembal-Saul, & Yoder, 2006).

Pedagogical discontentment is the measure of teachers’ satisfaction with their past and current teaching practice and is understood to be an indicator of whether an individual is poised for change (Southerland, Sowell, Blanchard, & Granger, 2011; Southerland, Sowell, & Enderle, 2011). Pedagogical discontentment embodies the dissonance required for learners (i.e., teachers) to seek out alternative

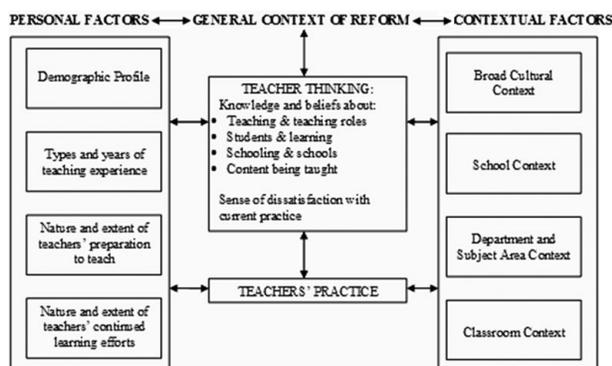


FIGURE 1. *The teacher-centered systemic reform framework (Gess-Newsome, Southerland, Johnston, & Woodbury, 2003). This model demonstrates the complex nature of influence that teacher thinking and professional development have on teachers' instructional practice.*

explanations or models for the phenomena of interest (i.e., science teaching). For PD to change teacher practice, pedagogical discontentment must be balanced with each teacher's self-efficacy as a teacher, because each must have a certain level of efficacy to believe that it is possible to participate in a practice. Strong teaching self-efficacy with elevated discontentment in practice is the combination that offers the best affective context to support change in practice (Southerland, Sowell, Blanchard, & Granger, 2011).

The final characteristic is a teacher's beliefs about science teaching and learning, which interact with and shape one's teaching practice. If researchers are to understand how a teacher interprets and practices inquiry, then these beliefs must be closely examined (Windschitl, 2002). The literature suggests that teachers' beliefs about teaching and their science teaching practices can be shaped toward more reform-minded instruction when PD experiences are structured to support such shifts (Herrington, Yezierski, Luxford, & Luxford, 2011; Silverstein et al., 2009) and when they are structured to influence a construct closely intertwined with beliefs that shape teachers' classroom practice (Enderle et al., 2014; Gess-Newsome et al., 2003; Gregoire, 2003). Furthermore, those beliefs interact in a complex manner with teachers' practice in their classrooms, influencing each other in recursive ways. Therefore, consideration of change in teachers' beliefs and practice as outcomes of PD should be coupled with exploration of more personal, affective states (Capps et al., 2012; Desimone, 2009).

#### *Previous Research Into the Influence of RET*

RET are a common form of teacher PD that has been offered in a host of settings nationwide for >35 years (Dubner et al., 2001). They are an approach to fostering teacher learning that privileges teacher engagement in the practices of a

discipline (NSF, 2016). Because of their long-term and prominent status in the PD portfolios of many university-school district partnerships, it seems prudent to explore the influence of such experiences on teachers' thinking and practice. Although RET occur in a variety of settings, most take place for 6 to 10 weeks in the summer. The goals of RET vary widely, with some seeking to increase teachers' comfort in scientific research or their knowledge of science and with others serving simply as a conduit for scientists to support the work of teachers. Although RET have a variety of goals, only those that have the goal of fostering teacher learning and change in practice are considered a form PD.

RET feature teacher immersion in an active, authentic research project at the "elbow" of a practicing scientist in a college or university, government, or company laboratory (Marx et al., 2004; Wade, Benson, & Switzer, 2012) to serve as a vehicle for teacher learning (NSF, 2014). Existing research provides evidence for the effectiveness of the RET programs on teachers' thinking (Dixon & Wilke, 2007; Schwartz, Westerlund, Garcia, & Taylor, 2010), practice (Blanchard, Southerland, & Granger, 2009; Herrington et al., 2011; Pop, Dixon, & Grove, 2010), and subsequent student achievement (Silverstein et al., 2009).

There is a growing body of literature that focuses on the influence of RET on change in teaching practice. As Sadler, Burgin, McKinney, and Ponjuan (2010) explain, the results of these studies are mixed. Blanchard et al. (2009) and Luft and Hewson (2014) suggest that changes in teachers' practice after participation in RET are determined by their incoming knowledge about teaching. Teachers who entered an RET program with "more sophisticated, theory-based understanding of teaching and learning" (Blanchard et al., 2009, p. 323) engage their students in the practices of science following the program. In contrast, teachers who did not hold a theory for student learning often explain that contextual barriers prevented their use of such practices. This research describes that RET may be more effective if the participants are prepared to learn from them, although it provides little insight into what specific features of the RET support teacher change.

Much of the extant research suggests that participation in RET does influence teachers' thinking—in terms of their self-efficacy and knowledge about the nature of science (Buck, 2003; Dresner & Worley, 2006; Grove, Dixon, & Pop, 2009; Miranda & Damico, 2013). However, current research is limited in terms of its success in teasing apart the specific features of teachers' thinking that are most influential in shaping the learning that occurs in RET; likewise, it does not identify the features of the RET that play the biggest role in the learning process. Research experiences have been treated as a "black box" by simply examining the influence that participation in research has on a selected outcome (e.g., beliefs, practice; see Figure 2). What is missing in this line of inquiry is an understanding of the essential features of these experiences—that

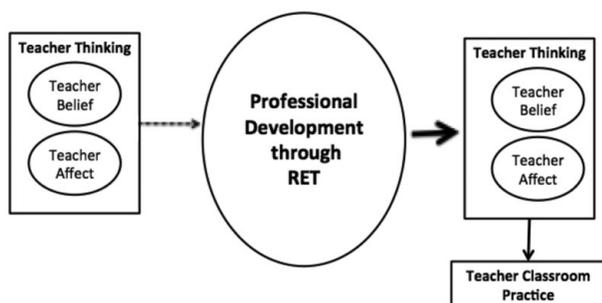


FIGURE 2. Influence of teacher thinking and professional development via Research Experiences for Teachers (RET) as drawn from the literature.

is, opening the black box to understand what aspects of the research experience influence teacher learning. We argue that given the importance of teachers' understanding of science as an epistemic activity, it is essential that we gain a more nuanced understanding of participation in these activities.

### Current Study

As a specific form of science teacher PD, RET can serve as potentially productive "field sites" to better understand the learning that results from teacher engagement in the disciplinary practices of science. We argue that RET programs can have a critical role in preparing teachers to enact the vision of science teaching and learning as described by the NGSS, by positioning them to engage in the disciplinary practices of science. Through a careful examination of the influence of these programs on teacher thinking and practice, this research can shed light on the nature of teacher learning in PD programs that immerse participants in disciplinary practices, as well as the essential aspects of PD that support that learning. Thus, the questions at the center of this research are as follows:

*Research Question 1:* In what ways does teachers' thinking—specifically, teaching self-efficacy, pedagogical discontentment, and beliefs about teaching—interact with research experience in an RET program to shape their practice?

*Research Question 2:* What are the features of the RET that are the most influential in teachers' learning, including change in their thinking and practice?

Due to the time frame of this study (2007–2012), both RET models examined in this study focused on scientific "inquiry" and how to teach through inquiry in the classroom. Significant for the purposes of this study, the Framework for K–12 Science Education in 2012 and the NGSS built upon it in 2013 have shifted the focus from "inquiry" to "scientific practices" to "better specify what is meant by inquiry in

science" (NRC, 2012, p. 30). The narrower conceptions of inquiry in many science classrooms mainly focused on developing and testing hypotheses based on data (Reiser, 2013). The eight practices of science in the NGSS are as follows:

- (1) asking questions and defining problems;
- (2) developing and using models;
- (3) planning and carrying out investigations;
- (4) analyzing and interpreting data;
- (5) using mathematics and computational thinking;
- (6) constructing explanations and designing solutions;
- (7) engaging in argument from evidence; and
- (8) obtaining, evaluating, and communicating information (NGSS Lead States, 2013).

As stated by Reiser (2013), these practices "incorporate much of what has been thought of as inquiry, but elaborate how to engage in the work of inquiry, and how this work is part of building knowledge" (p. 5). Moreover, these practices "represent the 'constituent elements' of scientific inquiry that have remained elusive as science educators have worked to foster inquiry-based science in K–12 classrooms" (Forbes, Biggers, & Zangori, 2013, p. 181). They also appear to be tightly aligned with the five essential features of inquiry that were set forth by the NRC in 2000, in which learners should (a) engage in scientifically oriented questions; (b) give priority to evidence; (c) formulate explanations from evidence to address scientifically oriented questions; (d) evaluate their explanations in light of alternative explanations, particularly those reflecting scientific understanding; and (e) communicate and justify their proposed explanations. As noted by Forbes et al. (2013), four of the practices (1, 3, 4, and 6) are in direct alignment with three of the essential features of inquiry (a, b, and c), whereas the other practices (2, 5, 7, and 8) involve multiple essential features of inquiry. There is agreement between the previous reform documents (NRC, 1996, 2000) and the current reform documents (NGSS Lead States, 2013; NRC, 2012) regarding vision of inquiry. Thus, although this study was designed and carried out prior to the new instantiation of science inquiry as practices, our analysis of the events within the PD applies to both these visions.

### Methods

To address the research questions, we examined change in teacher thinking and practice after teachers participated in two RET programs. This examination was conducted through a quantitative analysis of the change in thinking and practice as teachers participated in two models of RET: the Science Research (SciRes) program and the Science Pedagogy (SciPed) program.

### Participants

During 5 years of data collection (2007–2012), 106 teachers from urban, suburban, and rural schools participated in SciRes ( $n = 54$ ) or SciPed ( $n = 52$ )—two RET programs offered in a large university in the southeast. Both groups were diverse in terms of grade level taught, years of experience, geographic locations (11 states across the country), and the amount and nature of their teaching preparation and prior PD (Table 1).

### PD Programs

To identify the features of the RET that were most influential on teacher change, we compared teacher learning in the SciRes and SciPed models. Pertinent aspects of the two RET models are compared in Table 2.

*SciRes model.* The SciRes model, developed with funding from the National Science Foundation (supplement to DMR-0084173), was situated in a national laboratory. The goal of the 6-week summer SciRes program was to allow teachers to participate in authentic, engaging, ongoing scientific research in the laboratory of a scientist mentor. Research projects in the SciRes model focused on physical science concepts, such as materials research, superconductivity, and some geologic sciences and biomedical research. This model provided opportunities for teachers to learn about frontier science through engagement in research emerging from the work of the wider laboratory group and to rejuvenate their interest in science. The focus of this experience was on cutting-edge science without an extensive focus on translating teachers' learning into classroom practice. This model closely resembles the prevalent structure of RET nationwide.

*SciPed model.* The SciPed model, developed with funding from the NSF (ESI-9819431), was situated in a marine research station and had the goal of engaging teachers in scientific research and an in-depth reflective study of the learning that occurred during their research participation. This program was designed with the assumption that the research at the center of the RET should focus on teachers' questions that arise out of their experience in the research context (a marine environment) such that the research was personally relevant to the teachers. The SciPed model featured teachers' ongoing systematic reflection on how to structure such experiences to translate them into classroom practice. Thus, personal relevance of research and systematic pedagogical reflection are essential features of the SciPed model, making it a somewhat atypical RET program.

### Data Collection

As discussed above, due to the time frame of the study (2007–2012), both RET models attended to helping teachers

make sense of scientific “inquiry.” Given that focus, most of the methods and tools used in this study involve items and questions that specifically invoke the term *inquiry* as defined by its five essential features (NRC, 2000). However, the theoretical shift away from the concept of inquiry to the practices discussed above necessitated that our analysis of the events within the PD be analyzed in terms of this broader conceptualization of science learning.

Our research relied on a number of forms of data collected before, during, and after the RET experience: quantitative survey instruments, interview protocols, observations of research activities, and video recordings of classroom practice. The preprogram data (surveys, interviews, and video recordings of teaching) were collected from participants up to 5 weeks before the start of the RET. These same data were collected postprogram, 6 to 8 weeks after teachers returned to their classrooms.

*Surveys and interviews.* To determine teachers' state of preparation before the PD and as a way of measuring the influence of the PD, we measured several affective characteristics:

- (1) teachers' sense of their ability to teach science in general—that is, their general science teaching self-efficacy through the use of the Science Teaching Efficacy Belief Instrument (STEBI; Riggs & Enochs, 1990);
- (2) teachers' perceptions of their inquiry teaching abilities—namely, their science inquiry teaching self-efficacy through the Teaching Science as Inquiry Instrument (TSI; Smolleck et al., 2006);
- (3) teachers' fundamental beliefs about science teaching and learning—through the use of the Teacher Belief Inventory (TBI; Luft & Roehrig, 2007); and
- (4) teachers' state of dissatisfaction with their current teaching practices—specifically, their pedagogical discontentment through the Science Teachers' Pedagogical Discontentment Scale (Southerland et al., 2012).

For details about the survey instruments and interview protocols, see Table 3.

The TSI was developed around the five essential features of inquiry (NRC, 2000; Smolleck et al., 2006). An examination of the individual items on the TSI reveals that they map well onto the eight practices of science of the NGSS and that each of the eight practices are represented multiple times on the instrument. That said, note that the validity and reliability testing of the TSI was conducted around five clusters of items aligning with the five essential features of inquiry.

*Observations of teaching practice.* As a way to determine the degree to which teachers engaged in reform-based

TABLE 1

*Research Experiences for Teacher Participant Profile*

Program	Gender		Grade level <sup>a</sup>			Years of experience <sup>a</sup>			
	Female	Male	K–5	6–8	9–12	0–3	4–8	9–12	13–25
SciRes	33	21	18	16	14	17	16	5	12
SciPed	41	11	20	16	16	11	19	5	13

Note. SciRes = Science Research program; SciPed = Science Pedagogy program.

<sup>a</sup>Grade level and years of experience were undisclosed for a small number of teachers ( $n = 6$  and  $8$ , respectively).

teaching practices before and after the RET experience, teachers provided a video of a lesson (a minimum of at least 30 min showing the whole classroom perspective) that they considered to be their best example of inquiry teaching in their classrooms. These two videos—one before participation in the RET and one after it—provided researchers with observations of the teachers' classroom practice as related to their enactment of inquiry teaching involving the practices of science. Videos varied in length (e.g., 30–90 min) and lesson duration (e.g., from a single class period/block to multiple days).

To assess the degree to which the teaching practices captured on the video were reflective of the student-centered, collaborative, activity-driven instruction described in science education reforms (NRC, 2000)—which are necessary prerequisites for students to actually be engaged in the practices of science in the service of sense making—the videos were assessed through the Reformed Teacher Observation Protocol (RTOP; Sawada et al., 2002). The Likert-type scale of the RTOP ranged from 0 to 4 (e.g., from *never occurred* to *very descriptive*) and included 25 items divided into three subsets: lesson design and implementation (5 items), lesson content (10 items), and classroom culture (10 items). As such, the RTOP scores of teacher practices were used as an important measure of the outcome of participation in the PD.

*Observation of research.* For each RET model, observers detailed the day-to-day events and experiences of the teachers. Those field notes included the primary events of the day as scripted by program leaders, details of scientist-teacher interactions, activities of the day as enacted by each participant (along with specific time frames), and notes of weekly discussions with researchers and informal encounters among participants (lunch, bus trips, etc.). To triangulate these findings, teachers' work products (lesson plans, assignments, assessments) were also examined.

#### *Data Scoring*

*Surveys.* Survey instruments were scored through procedures described in the original publications. The results for each

were input into a data file; appropriate items were reverse scored; and totals for each instrument subscale were tallied.

*Interviews.* The analysis of TBI interview responses was guided by the rubrics developed by Luft and Roehrig (2007) examining beliefs along a continuum from traditional (teacher centered) to reform minded (student centered). Interview responses were assigned numeric scores according to categorical descriptions (1 = *traditional*, 2 = *instructional*, 3 = *transitional*, 4 = *responsive*, and 5 = *reform based*). Interview transcripts were analyzed by two members of the research team to determine interrater reliability (20% of the entire sample). The two raters' scores revealed substantial (Landis & Koch, 1977) interrater reliability ( $K = 0.794$ ,  $p < .001$ ; Pearson's  $r = .866$ ,  $p < .001$ ); thus, one rater separately analyzed the remaining transcripts.

*Observation of teaching practice.* Two researchers viewed and scored (0–100 points) each lesson independently using the RTOP rubric, and scores were discussed to determine reliability. When overall scores were separated by  $\geq 10$  points, the researchers discussed the video evidence for selecting particular scores and made adjustments based on agreement. These two scores were then averaged to determine an overall lesson score. When agreement was not possible, a third researcher reviewed the video, and the average of all three scores was taken. Having multiple raters over the course of several years limited the research team's ability to measure interrater reliability; however, strong correlation existed among raters (Pearson's  $r = .872$ ,  $p < .001$ ). The strong correlation among the three raters and the negotiation procedures enacted support the reliability of these data.

*Participation codes.* To determine the influences of distinct RET features, we identified specific structural variables of the research experiences (see Table 4 for details):

- the amount of social interaction,
- the primary intent of the research,
- the number of investigations completed,
- the type of teachers' products, and

TABLE 2  
*Comparison of Two Research Experiences for Teacher Models*

PD elements	SciRes	SciPed
Context	<ul style="list-style-type: none"> <li>• National High Magnetic Field Laboratory</li> </ul>	<ul style="list-style-type: none"> <li>• Coastal Biological Field Station</li> </ul>
Content knowledge	<ul style="list-style-type: none"> <li>• Exploration of science content through advanced scientific inquiries</li> </ul>	<ul style="list-style-type: none"> <li>• Exploration of marine biology content through participant-developed inquiries</li> </ul>
Social structure	<ul style="list-style-type: none"> <li>• One to two teachers placed within an ongoing research team</li> </ul>	<ul style="list-style-type: none"> <li>• Ten to 12 teachers working with one scientist (biologist) and a master teacher and organized into research groups of two to four based on their research question interests</li> </ul>
Authenticity of research	<ul style="list-style-type: none"> <li>• Investigations primarily focused on questions of current interest to the scientific community</li> </ul>	<ul style="list-style-type: none"> <li>• Investigations focused on scientific questions of interest to participants but whose answers might be known to the broader scientific community</li> </ul>
Intent of research	<ul style="list-style-type: none"> <li>• Investigations were designed to forward some aspect of the research teams activity</li> </ul>	<ul style="list-style-type: none"> <li>• Investigations were designed to teach participants about scientific concepts and practices</li> </ul>
Extended support	<ul style="list-style-type: none"> <li>• No formal activity structure</li> <li>• Some teachers participated in subsequent summers (not included in data collection again)</li> <li>• At the invitation of teachers; involved limited classroom outreach for activity demonstration by program director and/or staff</li> </ul>	<ul style="list-style-type: none"> <li>• No formal activity structure</li> <li>• At the invitations of teachers, involved minimal interaction with staff by a limited number of participants</li> </ul>
Modeling teaching	<ul style="list-style-type: none"> <li>• Sharing of inquiry lessons previously used by participants</li> </ul>	<ul style="list-style-type: none"> <li>• Practice implementation of newly developed inquiry lessons</li> </ul>
Reflection	<ul style="list-style-type: none"> <li>• Opportunities during afternoon whole group discussions</li> <li>• Maintenance of a laboratory notebook provided for recording individual thoughts and questions</li> </ul>	<ul style="list-style-type: none"> <li>• Multiple opportunities during whole group discussions and frequent journaling throughout stages of inquiry experiences with master teacher feedback</li> <li>• Daylong discussion activity with master teacher and scientist focused on modeling decisions</li> </ul>
Total time	<ul style="list-style-type: none"> <li>• Six-week summer program</li> <li>• “Teaching” topical focus for five 2-hour seminar sessions involving discussions and demonstrations by experts and teachers</li> </ul>	<ul style="list-style-type: none"> <li>• Six-week summer program</li> <li>• “Teaching” used as a key focus in iterative reflective whole group discussions and journaling dialectics between individual teachers and master teacher</li> </ul>

*Note.* SciRes = Science Research program; SciPed = Science Pedagogy program.

- engagement in each practice of science (asking questions, designing investigations, collecting and analyzing data, etc.; NGSS Lead States, 2013).

All field notes and work products relating to or produced by a teacher were coded through these variables. This resulted in a set of participation codes for each teacher.

#### *Statistical Data Analysis*

Only participants who completed both pre- and postprogram measures were used in the analysis of each data source. We explored the relationships among the various affective

measures, participation codes, research structure variables, and teachers’ beliefs and practice outcomes through confirmatory factor analysis (Harrington, 2009) and structural equation modeling (SEM) using the statistical program SPSS AMOS 22.0 (IBM, Chicago, IL). A two-phase approach was applied for SEM analysis (Anderson & Gerbing, 1988; Hair, Black, Babin, & Anderson, 2010). The measurement model was first estimated through confirmatory factor analysis to examine the overall fit, validity, and reliability of the model. Second, the SEM was adopted to examine hypotheses between constructs. We focused on pathway analysis to develop parsimonious models explaining the interrelationships among these multiple variables

TABLE 3  
*Descriptions of Survey Instruments and Interview Protocols*

Construct	Instrument	Description
Pedagogical discontentment	Science Teachers' Pedagogical Discontentment Scale	A 30-item multiple-choice Likert-scale instrument clustered around five subscales (Southerland et al., 2012) to measure teachers' discontentment with the current state of their classroom practices.
Self-efficacy in teaching	Science Teaching Efficacy Belief Instrument	A 23-question Likert-scale instrument designed to measure the level of self-efficacy and outcome expectancy of science teachers regarding their ability to teach science (Riggs & Enochs, 1990). The instrument has two subscales: (a) Personal Science Teaching Efficacy, which measures teachers' beliefs about their ability to teach science effectively; (b) Science Teaching Outcome Expectancy, which measures teachers' beliefs about whether their science teaching ability/skills will influence the outcome of learning.
Self-efficacy in teaching inquiry	Teaching Science as Inquiry Instrument	An instrument that includes 69 multiple-choice items designed to measure teachers' self-efficacy beliefs within the context of inquiry and reformed instruction (Smolleck, Zembal-Saul, & Yoder, 2006). The instrument has two subcategories: (a) Personal Science Teaching Efficacy, which contains 34 items clustered around five component factors; (b) Science Teaching Outcome Expectancy, developed around the essential features of classroom inquiry as described by the National Research Council (2000).
Teacher beliefs	Teacher Belief Inventory	A seven-item semistructured interview protocol (Luft & Roehrig, 2007) used to elicit teaching and learning beliefs in science and mathematics.

TABLE 4  
*Structural Variables of the Research Experience*

Label	Description	Values
SI	Social interaction during research—describes the amount of social interaction that a teacher experienced during the majority of his or her research.	0 = pure observation, 1 = independent work, 2 = scheduled interaction, 3 = side-by-side laboratory immersion
INT	Intent of the research—dichotomous variable indicating whether teacher's research was intended to develop science knowledge or personal understanding.	0 = authentic to science, 1 = authentic to the teacher
QR	Quantity of research projects—total number of research projects in which the teacher participated.	Descriptive count
CP	Culminating product—dichotomous variable describing the nature of the teacher's RET capstone.	0 = focus on research project, 1 = focus on teaching of science
SciPrac	Scientific practices enacted by teachers—describes engagement in each of eight NGSS scientific practices.	0 = no teacher engagement, 1 = teacher engagement

Note. RET = Research Experiences for Teachers; NGSS = Next Generation Science Standards.

(Kline, 2011). Multiple models were evaluated by means of common fit indices estimating how well they fit the data, including root mean square error of approximation (Steiger, 1990) and confirmatory fit index (Bentler, 1990). During the analysis, several variables were found to have high collinearity with the program variable (SciRes or SciPed), meaning that the variation in those measures was too heavily correlated with a specific RET program to be useful in constructing a more general model for these experiences, and they were not used in the final models. Several models comprising different variable combinations were evaluated by means of common indices for describing how well a proposed

model fit the actual data. The assessment of factor loadings, convergent validity, discriminant validity, and reliability was performed for the latent constructs through a confirmatory factor analysis (Hair et al., 2010).

## Results

The measures for establishing validity and reliability, which are shown in Table 5, indicate that the measures used satisfy the suggested thresholds, except that the average variance extracted (AVE) for preprogram STEBI–Personal Science Teaching Efficacy subscale (STEBIP) was <0.50. A

TABLE 5  
Validity and Reliability of Measurements

	CR	AVE	MSV	ASV	TSIP	STPD	STEBIP
TSIP	0.924	0.710	0.082	0.058	0.843		
STPD	0.879	0.595	0.085	0.060	-0.186	0.771	
STEBIP	0.793	0.251	0.085	0.084	0.287	-0.292	0.501

Note. CR, composite reliability; AVE, average variance extracted; MSV, maximum shared variance; ASV, average shared variance; TSIP, Teaching Science as Inquiry Instrument–Personal Science Teaching Efficacy subscale; STPD, Science Teachers’ Pedagogical Discontentment Scale; STEBIP, Science Teaching Efficacy Belief Instrument–Personal Science Teaching Efficacy subscale.

TABLE 6  
Goodness-of-Fit Indices for the Structural Equation Model

Fit index	Criterion	Result/value
$\chi^2 / df$	<3	1.57
Model <i>p</i> value	>.05	.10
RMSEA	<0.08	0.077
PCLOSE	>.05	.24
CFI	>0.8	0.87

Note. RMSEA = root mean square error of approximation; PCLOSE = *p* value of close fit; CFI = comparative fit index.

potential reason for the low AVE for STEBIP was the relatively small sample size in our study. Note, however, that the STEBI is a scale that has been frequently used and well validated in the science-education field (Enoch & Riggs, 1990), justifying inclusion of its subscale (STEBIP) as a measured variable in the SEM model.

Assessment of the adequacy of the structural model (AMOS 22.0) confirmed the extent to which the relationships specified by the model were consistent. Fit indices were used to examine model fit based on the suggestions of Hair et al. (2010) and Kline (2011)—namely, chi-square per degree-of-freedom ratio ( $\chi^2/df$ ), root mean square error of approximation, and comparative fit index were used. These indicators, with their values and verification criteria as reported in Table 6, revealed an acceptable structural equation model. The overall fit of the model to the data was good, with a nonsignificant chi-square test ( $\chi^2 = 17.28, df = 11, p = .10, \chi^2/df < 3$ ) and with other fit indices within acceptable ranges (confirmatory fit index = 0.87, root mean square error of approximation = 0.077).

Results of the analysis for the structural model are presented in Figure 3. The estimated path coefficient (standardized) is specified on each link. The  $R^2$  statistic is indicated above each endogenous construct. Table 7 summarizes the direct, indirect, and total effects. Significant regression relationships exist between many of the variables present in the model and are represented numerically in Figure 3. As a measured variable increases by one standardized unit, the

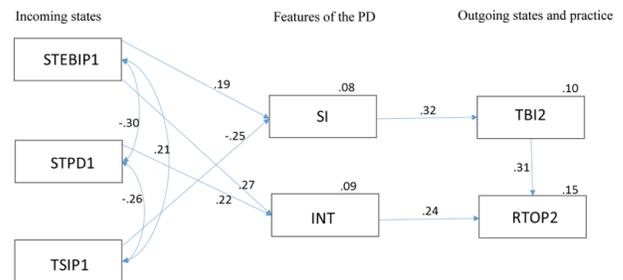


FIGURE 3. Structural equation model for interactions of affect, beliefs, and practice as influenced by participation in features of Research Experiences for Teachers. As a measured variable increases by one standardized unit, the variable to which it is connected changes by the value above the arrow. STEBIP = Science Teaching Efficacy Belief Instrument–Personal Science Teaching Efficacy subscale (a measure of teachers’ thoughts about their ability to teach science effectively); STPD = Science Teachers’ Pedagogical Discontentment Scale; TSIP = Teaching Science as Inquiry Instrument–Personal Science Teaching Efficacy subscale; SI = social interactions; INT = intent of the research; TBI = Teacher Belief Inventory; RTOP = Reformed Teacher Observation Protocol. For each, “1” denotes preprogram measure, and “2” denotes postprogram measure.

variable to which it is connected changes by the value above the arrow—thus,

- for a one-unit increase in program entry scores (preprogram data) on the teachers’ science inquiry teaching self-efficacy, the amount of social interaction in teachers’ RET research decreased by 0.251 units;
- for a one-unit increase in general science teaching self-efficacy, the intent of their RET research increased by 0.271 units;
- for a one-unit increase in pedagogical discontentment, the intent of their RET research increased by 0.222 units;
- for a one-unit increase in the social interaction in teachers’ RET research, TBI scores following the RET (postprogram data) increased by 0.322 units;

TABLE 7

*Standardized Total Effects of Relationships Between Model Variables*

	STEBIP <sup>a</sup>	TSIP <sup>a</sup>	STPD <sup>a</sup>	SI	TBI <sup>b</sup>	INT	Total, $R^2$
SI	0.19 <sup>c</sup>	-0.25 <sup>c</sup>	—	—	—	—	0.08
TBI <sup>b</sup>	0.06 <sup>d</sup>	-0.08 <sup>d</sup>	—	0.32 <sup>c</sup>	—	—	0.10
INT	0.27 <sup>c</sup>	—	0.22 <sup>c</sup>	—	—	—	0.09
RTOP <sup>b</sup>	0.08 <sup>d</sup>	-0.03 <sup>d</sup>	0.05 <sup>d</sup>	0.10 <sup>d</sup>	0.31 <sup>c</sup>	0.24 <sup>c</sup>	0.15

*Note.* See Figure 3. STEBIP, Science Teaching Efficacy Belief Instrument; TSIP, Teaching Science as Inquiry Instrument–Personal Science Teaching Efficacy subscale; STPD, Science Teachers’ Pedagogical Discontentment Scale; SI, social interactions; TBI, Teacher Belief Inventory; INT, intent of research; RTOP, Reformed Teacher Observation Protocol.

<sup>a</sup>Preprogram measure.

<sup>b</sup>Postprogram measure.

<sup>c</sup>Direct effect (as column variable increases by one unit, row variable changes by listed value).

<sup>d</sup>Indirect effect of column variable on row variable.

- for a one-unit increase in the intent of the teachers’ RET research, their reformed teaching scores following the RET increased by 0.240 units; and
- for a one-unit increase in the TBI scores following the RET, their reformed teaching scores following the RET increased by 0.306 units.

The results indicate that the model explained 10.4% of the variance in teacher belief and 15.3% of the variance in reformed teaching.

#### *Examining Influence of Incoming States*

The SEM highlights the importance of teachers’ incoming affective states, which were found statistically distinct but related constructs (Figure 3). Teachers’ incoming pedagogical discontentment was negatively correlated with their confidence in their personal ability to teach science in general and through inquiry as defined by the TSI specifically. That is, teachers who were concerned about their past success in teaching science (i.e., higher pedagogical discontentment) were also concerned about their future ability to teach science in general or through inquiry (lower self-efficacy).

Note that participation in specific features of a RET were at the discretion of the teachers, as they selected the RET program to which they would apply and the sort of research in which they wanted to be engaged. Thus, the model suggests that incoming affective states influenced teacher selection of RET in terms of social interaction. Teachers with more confidence in their ability to teach through inquiry selected RET that involved less social interaction—which were more prominent at the SciRes program. Teachers with stronger general science teaching self-efficacy but lower inquiry-teaching self-efficacy selected more socially interactive, group-centered research contexts—which were more prominent at the SciPed program. Overall, the constructs of science teaching self-efficacy and teachers’ self-efficacy for inquiry accounted for 8% of the variance in social interaction.

As shown in Figure 3, incoming affective states also seemed to influence teachers’ choices of RET based on the intent of the research. The SEM indicates that teachers with a higher incoming general science teaching self-efficacy were more likely to select research whose design was intended to directly contribute to their personal understanding of science and science teaching, as opposed to selecting cutting-edge science research directed at developing new knowledge for the discipline. Likewise, teachers with a stronger sense of pedagogical discontentment selected research that was designed for personal relevance. Overall, the constructs of science teaching self-efficacy and pedagogical discontentment accounted for 9% of the variance in the intent of the research.

#### *Examining the Influence of RET Outcomes*

The SEM also suggests that certain PD features shaped the teachers’ outcomes, in terms of both thinking and practice. The amount of sustained social interaction during research, with fellow teachers and laboratory members, had a positive influence on teacher beliefs about science teaching (Figure 3) as examined with the TBI (Luft & Roehrig, 2007). That is, more opportunities to engage with others in sense making surrounding the research shifted teaching beliefs toward more student-centered perspectives. Social interactions during research allowed participants to discuss and evaluate their experiences with peers, enriching their knowledge building about science, science practices, and science teaching. Such interactions were more prevalent for teachers participating in the SciPed program than for teachers in the SciRes program. Of all the features of the RET examined (see Table 4), the social nature of the research experiences was the only one to play an important role in shaping teachers’ thinking—specifically, their beliefs about teaching. Indeed, of all the features, social interaction during research had the largest direct effect on shifting teachers’ beliefs about science teaching and learning. Overall, this

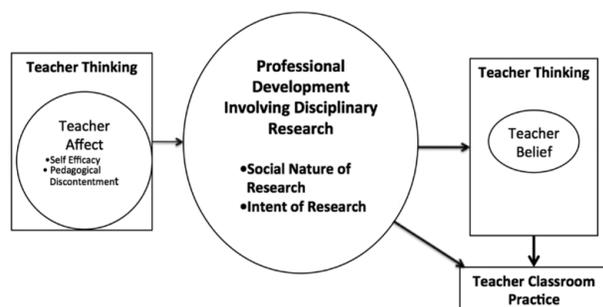


FIGURE 4. *Interaction of teacher learning and disciplinary research-based professional development drawn from the findings of the present study.*

aspect of the model accounted for 10.4% of the variance in teachers' beliefs.

Finally, the SEM suggests that teachers' classroom practice (as measured by the RTOP scores of teaching videos) was influenced through two pathways (Figure 3). First, extended opportunities to focus on research intended to develop their personal understandings about science and its practices had a direct positive impact on practice. It is important to note that of all the features of the RET examined (see Table 4), the intent of the research was the only one to play an important and direct role in shaping teachers' classroom practices—that is, research that was designed to enhance teacher learning, as opposed to research designed to forward the body of knowledge of science as a whole, had a direct influence on teachers' classroom practice. The other features of the RET experience—the social interactions, the quantity of research, the culminating process, or the scientific practices in which teachers had an opportunity to engage—did not play such a direct role. Second, the model suggests that teachers' classroom practices were enhanced indirectly through shifting their beliefs about teaching, which in turn shaped their classroom practice. Of the two routes of influence, shifting teachers' beliefs had the largest effect on practice. Overall, the model accounted for 15.3% of the variance in teachers' classroom practice after RET experiences.

## Discussion

There is a new emphasis on participation in the disciplinary practices of science to engage students in productive sense making about the natural world as a vehicle to support their science learning and as an essential feature of broader science proficiency (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010; NRC, 2014; Zwiers, O'Hara, & Prichard, 2014). It has been widely acknowledged that the kinds of instruction and learning opportunities to be engendered in such efforts are often in stark contrast to normal instruction—a situation that will call for effective vehicles of teacher PD (Gulamhussein,

2013; Walters, Scheopner Torres, Smith, & Ford, 2014; Wilson, 2013). While numerous options for programs have been designed for teacher learning in science, the field is just now focusing on how such experiences influence teacher change and what essential features of PD influence that change. Mindful of these gaps in the literature, we designed this research to examine PD centered on the engagement of teachers in the disciplinary practices of science.

Figure 4 represents a model of the interaction of teacher learning and disciplinary research-based PD drawn from the findings of the present study. Our findings support much of the model developed from a review of the literature—that is, teacher thinking influences a teacher's engagement in PD, which then can influence both teacher thinking and classroom practice. However, our findings suggest some refinements of this general model when the PD examined is an RET program.

### *Teacher Thinking*

The aspect of teacher thinking that explained more of the variance in teacher learning was affect—in the form of personal self-efficacy and science teacher pedagogical discontentment. Teachers' affective states (e.g., self-efficacy) played an important role in shaping their learning (e.g., Smolleck et al., 2006; Southerland, Sowell, Blanchard, & Granger, 2011). The SEM (Figure 3) suggests an important interaction between incoming affective states and the type of PD that teachers seek. Teachers with higher science inquiry teaching self-efficacy were more likely to select research experiences that focused on solitary, less social efforts than were their colleagues who were less confident in their use of scientific practices related to inquiry, who gravitated to research experiences that were more socially interactive. Likewise, teachers who were discontented with their past teaching efforts but convinced that they were effective science teachers were more likely to select research that was for their personal understanding, as opposed to gravitating to more cutting-edge science research opportunities. This finding supports the importance of considering teachers' incoming affective states before entering a PD experience. As affect influences the types of PD that teachers seek out, it is reasonable to consider that this influence might extend its impact to the willingness of participants to consider and internalize the messages and models of teaching being emphasized during the experience (Ebert & Crippen, 2010; Gregoire, 2003).

### *Important Features of PD That Engages Teachers in Research*

Selection of research experience is also important, as the SEM (Figure 3) suggests that research opportunities with sustained social interactions are more likely to influence teachers' beliefs about their teaching, shifting their beliefs

toward the importance of student participation in the practices of science. These social interactions refer to those with peers and laboratory group members, focused on the research. Of all the features of the research experience examined—social interactions, intent of the research, quality of research project, culminating product, and scientific practices—the nature of the social interaction showed the largest influence on teachers' beliefs, demonstrating an important pathway for shaping classroom practices. The SEM quantitatively reflects the important interaction between the instructional practices that teachers implement in their classrooms and their beliefs about the teaching and learning of science, as established in the literature (Crawford, 2007; Cross & Hong, 2012).

Furthermore, the model highlights an important but unexpected pathway for the role of research in directly shaping practice: For teachers who elected to participate in research directed toward their personal understanding of science and its practices, there was a direct pathway to changing their teaching practices, a pathway that did not involve mediation by a change in teacher thinking. This same direct relationship was not seen for teachers who opted to engage in research identified as cutting-edge science—suggesting that such experiences are not as powerful a force for informing a teachers' classroom work. Indeed, this model suggests that personally relevant research experiences—that is, research that arises from teachers' own questions about phenomena—were the more powerful force. This problem of translating frontier science research experiences into productive science pedagogy has been observed in other contexts involving teachers engaged in research (Blanchard et al., 2009; Jeanpierre, Oberhauser, & Freeman, 2005).

Our findings reflect the importance of teacher thinking both prior to and following a research experience in PD as identified in the conceptual model shown in Figure 4. Teacher thinking encompasses several distinct elements, with beliefs and affect being unique constructs that do not share a quantitative correlative relationship in the model; however, the two self-efficacy constructs (i.e., self-efficacy in teaching and self-efficacy in teaching inquiry) do manifest significant positive relationships with each other and inverse relationships with the affective construct of pedagogical discontentment (see Figure 3). All of these are seen as quantitative measures of theoretically aligned relationships. Interestingly, an indirect relationship exists between teacher beliefs and affect, where self-efficacy and discontentment interact with how the research experience is processed socially, which ultimately operates to shift beliefs about teaching and learning. Also, teacher affect can mediate changes to teaching practice, positively and directly, through research experiences that have relevance for personal beliefs about science and its practices. These specific theoretical insights resonate with broader models for teachers' cognition that identify affective constructs as critically shaping

teachers' initial assessment of ideas as a threat or a challenge (Gregoire, 2003). As teachers originally assess a new model of instruction, their beliefs about teaching and learning serve as filters that they use to judge the coherency and plausibility of the new approach (Jones & Carter, 2007; Pajares, 1997). These theoretical relationships in combination with the results of this study provide further support for simultaneously focusing on teacher beliefs and affect when investigating the impact of teacher PD on teacher learning and subsequent classroom practice.

### *Research Experiences as PD*

Although links among PD, teacher beliefs, and teaching practice have been acknowledged in the literature, much of this research has been theoretical or based in case study (e.g., Gess-Newsome et al., 2003; Hughes, Molyneaux, & Dixon, 2012; Luft & Hewson, 2014). This investigation provides important quantitative support for the direct influence that PD through research experiences can have on shaping practices, as well as for indirect influences on shaping beliefs that in turn shape practice.

These findings suggest that if research experiences are to be effective forms of PD, then one of the critical features of these experiences is that the research should be conducted in a social context, as part of a group or research team. Here, it is important to highlight that by “socially” we are suggesting that teachers have a community with which to make sense of their research experiences, an idea introduced in previous smaller-scale, in-depth studies (Hughes et al., 2012; Pop et al., 2010; Varelas, House, & Wenzel, 2005). This finding expands the scope of collaboration emphasized across PD frameworks (Desimone, 2009; Luft & Hewson, 2014; Wilson, 2013) to include communities with members who may not necessarily be teachers but knowledgeable others with respect to the focal activity of the experience. Although broadly present in earlier frameworks for effective PD, this particular characteristic of the research experience was not expected to be a significant mediator in the research experience. However, the Framework for K–12 Science Education (NRC, 2012) and NGSS (NGSS Lead States, 2013) emphasize that science learning requires engagement in the epistemic practices of science. “Productive epistemic discourse” involves interactions that allow learners to create meaning and understanding through constructing explanations, engaging in argument from evidence, and evaluating and communicating information. Science learning, as described by the NGSS, requires students to engage in productive epistemic discourse with others—work that involves talk, joint attention, and shared activity aimed at the construction and critique of ideas (Ford, 2008). Viewed from this lens, the import of social interactions in shaping teachers learning from research experience is clear. Given this finding, we have revised the original conceptual framework that informed the

structure of this study to also include social interaction with peers and/or knowledgeable others as another critical aspect of PD around disciplinary practices (see Figure 4).

Our findings further suggest that if a research experience is intended to be a cornerstone of a PD experience, that research should have personal relevance to the teachers' understanding of science and its relation to their work in classrooms. The immersion in and modeling of research during PD, as highlighted by a number of reviews of effective science PD (Capps et al., 2012; Wilson, 2013), should incorporate the specific aspects identified in this study if the research is to engender changes in teachers' classroom practice.

### **Implications, Limitations, and Further Research**

Our results suggest that research participation in itself is not sufficient to shape teachers' use of scientific practices in the classroom, although carefully crafted research experiences can do so. Teachers must have an opportunity to make sense of their research experiences with others if they are to undergo the necessary changes in affect and belief to result in changes in practice. Indeed, while the archetypes of a "lone researcher" or even of a novice working with an uncommunicative scientist may mesh nicely with a common conception of the work of scientists, they result in very few changes in teachers' beliefs or practice. Instead, research groups in which teachers had a chance to make sense of their experiences with other teachers and with other scientists were essential to learning.

Likewise, our results confront the common quest for the authenticity of educational programs (Barab & Hay, 2001). Many of the teachers in our study were assisting cutting-edge science research under the guidance of scientists working in one of the foremost national laboratories. Certainly, many would consider work of this nature "authentic." In contrast, other teachers in the RET worked on projects that were not new to science but instead new to the teacher and arose out of their own attempts to understand phenomena, and so their research was personally relevant to them. Our work suggests that teachers engaged in research driven by their own questions were the ones who demonstrated direct gains in their use of disciplinary practices in their classrooms. In these personally relevant experiences, teachers were in charge of not only developing the questions to be pursued but also designing and conducting the investigations. One can certainly argue that such experiences are "authentic" to the teacher, if not to the larger field of science. Engagement in research projects intended to have personal relevance had a direct influence on aspects of affect, which in turn would position teachers to benefit from future PD experiences. As seen in the work of Buxton (2006), in the design of productive learning experiences, it is important to attend to the issue of personal relevance and engagement.

In terms of research on teacher learning, our findings suggest that it is essential to capture teachers' affective states at the outset of a PD experience, as these states serve to shape the influence of the PD experience. In addition, there is significance in that at least two pathways of a research PD experience influence teachers' practice: indirectly, through beliefs as previously identified, and directly, through the influence of the intent of the research. These findings are particularly important for shaping theories of teacher learning.

Finally, our findings suggest that no single template exists for an effective PD. Incoming affective states shape the sorts of experiences that teachers will embrace. Like their own students, teachers are not "blank slates" as they enter PD. Instead, consideration of their prior knowledge, beliefs, habits, and dispositions toward new ideas is essential to facilitating effective PD.

It is important to acknowledge the limitations of this work. While >100 teachers were included in the study, they worked in variety of contexts and grade levels; they were drawn from across the nation; and each teacher self-selected to participate in the RET. Because of the reliance on volunteers, there was a marked gender imbalance among participants (with women outnumbering men), as well a smaller number of teachers with 9 to 12 years of experience. Although our analysis failed to find an influence of gender or experience on the relationships documented, this omission may be due to the relatively small sample sizes for men and for teachers with 9 to 12 years of experience. While this unequal distribution may be reflective of the demographics of science teachers at large, it does call into question the generalizability of the findings. Indeed, our findings should be generalized only to teachers who select to participate in an RET program, which, because of its intensive nature (6 weeks in the summer), may omit a wide swath of teachers.

A second limitation of this work comes from the research design. While one might posit the need for an eventual randomized controlled experimental design to allow for a determination of causality, the focus on the preparedness of teachers as playing a central role in their learning from RET suggests that it is necessary to employ volunteers, because their readiness to learn from the RET is a fundamentally important influence on their learning. However, future research should be designed to select a more equal distribution of gender and experience, with a larger sample size, to replicate the current study and to check the fitness of the models given the same measures.

### **Funding**

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was funded by the National Science Foundation, DISCOVERY RESEARCH K-12 Program, Award Number #0553769.

## References

- Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, *103*(3), 411–423.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.
- Banilower, E. R., Smith, P. S., Weiss, I. R., Malzahn, K. A., Campbell, K. M., & Weis, A. M. (2013). *Report of the 2012 National Survey of Science and Mathematics Education*. Chapel Hill, NC: Horizon Research.
- Barab, S. A., & Hay, K. E. (2001). Doing science as the elbows of experts: Issues related to the science apprenticeship camp. *Journal of Research in Science Teaching*, *38*(1), 70–102.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*(2), 238–246.
- Blanchard, M. R., Southerland, S. A., & Granger, 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.
- Buck, P. (2003). Authentic research experiences for Nevada high school teachers and students. *Journal of Geoscience Education*, *51*(1), 48–53.
- Buxton, C. A. (2006). Creating contextually authentic science in a “low-performing” urban elementary school. *Journal of Research in Science Teaching*, *43*, 695–721.
- Capps, D. K., Crawford, B. A., & Constan, M. A. (2012). A review of empirical literature on inquiry professional development: Alignment with best practices and a critique of the findings. *Journal of Science Teacher Education*, *23*(3), 291–318.
- Crawford, B. A. (2000). Embracing the essence of inquiry: New roles for science teachers. *Journal of Research in Science Teaching*, *37*(9), 916–937.
- Crawford, B. A. (2007). Learning to teach science as inquiry in the rough and tumble of practice. *Journal of Research in Science Teaching*, *44*(4), 613–642.
- Cross, D. I., & Hong, J. Y. (2012). An ecological examination of teachers’ emotions in the school context. *Teaching and Teacher Education*, *28*(7), 957–967.
- Desimone, L. (2009). Improving impact studies of teachers’ professional development: Toward better conceptualizations and measures. *Educational Researcher*, *38*(3), 181–199.
- Desimone, L. M. (2011). A primer on effective professional development. *Phi Delta Kappan*, *92*(6), 68–71.
- Dixon, P., & Wilke, R. (2007). The influence of a teacher research experience on elementary teachers’ thinking and instruction. *Journal of Elementary Science Education*, *19*(1), 25–43.
- Dresner, M., & Worley, E. (2006). Teacher research experiences, partnerships with scientists, and teacher networks sustaining factors from professional development. *Journal of Science Teacher Education*, *17*(1), 1–14.
- Dubner, J., Silverstein, S., Carey, N., Frechtling, J., Busch-Johnsen, T., Han, J., . . . Zounar, E. (2001). Evaluating science research experience for teachers programs and their effects on student interest and academic performance: A preliminary report of an ongoing collaborative study by eight programs. *Materials Research Science Proceedings*, *684*. doi:http://dx.doi.org/10.1557/PROC-684-GG3.6
- Ebert, E. K., & Crippen, K. J. (2010). Applying a cognitive-affective model of conceptual change to professional development. *Journal of Science Teacher Education*, *21*(3), 371–388.
- Enderle, P., Denzau, M., Roseler, K., Southerland, S. A., Granger, E., & Hughes, R. (2014). Examining the influence of RET’s on science teachers’ beliefs and practice. *Science Education*, *98*, 1077–1108.
- Enoch, L. G., & Riggs, L. M. (1990). Further development of an elementary science teaching efficacy belief instrument: A pre-service elementary scale. *School Science and Mathematics*, *90*(8), 694–706.
- Forbes, C. T., Biggers, M., & Zangori, L. (2013). Investigating essential characteristics of scientific practices in elementary science learning environments: The Practices of Science Observation Protocol (P-SOP). *School Science and Mathematics*, *113*, 180–190.
- Ford, M. J. (2008). Disciplinary authority and accountability in scientific practice and learning. *Science Education*, *92*(3), 404–423.
- 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.
- Granger, E. M., Bevis, T., Saka, Y., Southerland, S. A., & Sampson, V. (2012). Efficacy of student centered instruction in supporting student science learning. *Science*, *338*(6103), 105–108.
- Gregoire, M. (2003). Is it a challenge or a threat? A dual-process model of teachers’ cognition and appraisal processes during conceptual change. *Educational Psychology Review*, *15*, 147–179.
- Grove, C. M., Dixon, P. J., & Pop, M. M. (2009). Research experiences for teachers: Influences related to expectancy and value of changes to practice in the American classroom. *Professional Development in Education*, *35*(2), 247–260.
- Gulamhussein, A. (2013). *Teaching the teacher: Effective professional development in an era of high stakes accountability*. Alexandria, VA: National School Boards Association.
- Hair, J., Black, W., Babin, B., & Anderson, R. (2010). *Multivariate data analysis* (7th ed.). Upper Saddle River, NJ: Prentice Hall.
- Harrington, D. (2009). *Confirmatory factor analysis*. New York, NY: Oxford University Press.
- Herrington, D. G., Yezierski, E. J., Luxford, K. M., & Luxford, C. J. (2011). Target inquiry: Changing chemistry high school teachers’ beliefs about inquiry instruction and their classroom practice. *Chemical Education Research and Practice*, *12*(1), 74–84.
- Hughes, R., Molyneaux, K., & Dixon, P. (2012). The role of scientist mentors on teachers’ perceptions of the community of science during a summer research experience. *Research in Science Education Journal*, *42*(5), 915–941.
- Jeanpierre, B., Oberhauser, K., & Freeman, C. (2005). Characteristics of professional development that effect change in secondary science teachers’ classroom practices. *Journal of Research in Science Teaching*, *42*(6), 668–690.
- Jones, G., & Carter, G. (2007). Science teacher attitudes and beliefs. In N. Lederman & S. Abell (Eds.), *Handbook of research on science education* (pp. 1067–1104). New York, NY: Routledge.

- Kline, B. B. (2011). *Principles and practice of structural equation modeling*. New York, NY: Guilford Press.
- Krajcik, J., Codere, S., Dahsah, C., Bayer, R., & Mun, K. (2014). Planning instruction to meet the intent of the Next Generation Science Standards. *Journal of Science Teacher Education*, 25(2), 157–175.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159–174.
- Little, J. W. (1993). Teachers' professional development in a climate of educational reform. *Educational Evaluation and Policy Analysis*, 15(2), 129–151.
- Luft, J. A., & Hewson, P. (2014). Research on teacher professional development in science. In N. G. Lederman & S. K. Abell (Eds.), *Handbook of research in science education* (2nd ed.). New York, NY: Routledge.
- Luft, J. A., & Roehrig, G. H. (2007). Capturing science teachers' epistemological beliefs: The development of the teacher beliefs interview. *Electronic Journal of Science Education*, 11(2), 38–63.
- 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.
- Miranda, R. J., & Damico, J. B. (2013). Science teachers' beliefs about the influence of their summer research experiences on their pedagogical practices. *Journal of Science Teacher Education*, 24(8) 1241–1261.
- Moon, J., Michaels, S., & Reiser, B. (2012, November). Science standards require a teacher-learning rethink. *Education Week*, 32(13).
- National Academies of Sciences, Engineering, and Medicine. (2015). *Science teachers learning: Enhancing opportunities, creating supportive contexts*. Washington, DC: National Academies Press.
- National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards*. Washington, DC: Authors.
- National Research Council. (1996). *National Science Education Standards*. Washington, DC: National Academies Press.
- National Research Council. (2000). *Inquiry and National Science Education Standards: A guide for teaching and learning*. Washington, DC: National Academies Press.
- National Research Council. (2012). *A framework for K–12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: National Academies Press.
- National Research Council. (2014). *Literacy for science: Exploring the intersection of the Next Generation Science Standards and Common Core for ELA standards*. Washington, DC: National Academies Press.
- National Science Foundation. (2016). *Research experiences for teachers (RETs) in engineering and computer science*. Retrieved from [https://www.nsf.gov/funding/pgm\\_summ.jsp?pims\\_id=505170](https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505170)
- NGSS Lead States. (2013). *Next Generation Science Standards: For states, by states*. Washington, DC: National Academies Press.
- Pajares, F. (1997). Current directions in self-efficacy research. In M. Maehr & P. R. Pintrich (Eds.), *Advances in motivation and achievement* (pp. 1–49). Greenwich, CT: JAI Press.
- Passmore, C. (2014, November 10). Implementing the Next Generation Science Standards: How your classroom is framed is as important as what you do in it. *NSTA Blog*.
- Pop, M. M., Dixon, P., & Grove, C. (2010). Research experiences for teachers (RET): Motivation, expectation, and changes to teaching practices due to professional program involvement. *Journal of Science Teacher Education*, 21, 127–147.
- Reiser, B. J. (2013, September). *What professional development strategies are needed for successful implementation of the Next Generation Science Standards*. Paper presented at the Invitational Research Symposium on Science Assessment, Washington, DC.
- Riggs, I. M., & Enochs, L. E. (1990). Toward the development of an elementary teacher's science teaching efficacy belief instrument. *Science Education*, 74, 625–637.
- Roehrig, G. H., & Kruse, R. A. (2005). The role of teachers' beliefs and knowledge in the adoption of a reform-based curriculum. *School Science and Mathematics*, 105(8), 412–422.
- Sadler, T. D., Burgin, S., McKinney, L., & Ponjuan, L. (2010). Learning science through research apprenticeships: A critical review of the literature. *Journal of Research in Science Teaching*, 47(3), 235–256.
- Sawada, D., Piburn, M., Judson, E., Turley, J., Falconer, K., Benford, R., & Bloom, I. (2002). Measuring reform practices in science and mathematics classrooms: The Reformed Teaching Observation Protocol. *School Science and Mathematics*, 102(6), 245–253.
- Schwartz, R. S., Westerlund, J. F., Garcia, D. M., & Taylor, T. A. (2010). The impact of full immersion scientific research experiences on teachers' views of the nature of science. *Electronic Journal of Science Education*, 14(1). Retrieved from <http://ejse.southwestern.edu/article/view/7325/5623>
- Silverstein, S. C., Dubner, J., Miller, J., Glied, S., & Loike, J. D. (2009). Teachers' participation in research programs improves their students' achievement in science. *Science*, 326(5951), 440–442.
- Smolleck, L., Zembal-Saul, C., & Yoder, E. (2006). The development and validation of an instrument to measure preservice teachers' self-efficacy in regard to the teaching of science as inquiry. *Journal of Science Teacher Education*, 17(2), 137–163.
- Southerland, S. A., Nadelson, L., Sowell, S., Kahveci, M., Saka, Y., & Granger, E. M. (2012). Measuring one aspect of teachers' affective states: Development of the science teachers' pedagogical discontentment scale. *School Science & Mathematics*, 112(8), 483–494.
- Southerland, S. A., Sowell, S., Blanchard, M., & Granger, D. E. (2011). Exploring the construct of pedagogical discontentment: A tool to understand science teachers' openness to reform. *Research in Science Education*, 41(3), 299–319.
- Southerland, S. A., Sowell, S., & Enderle, P. (2011). Science teachers' pedagogical discontentment: Its sources and potential for change. *Journal of Science Teacher Education*, 22(5), 437–457.
- Steiger, J. H. (1990). Structural model evaluation and modification. *Multivariate Behavioral Research*, 25, 214–212.
- Varelas, M., House, R., & Wenzel, S. (2005). Beginning teachers immersed into science: Scientist and science teacher identities. *Science Education*, 89(3), 492–516.

- Wade, C., Benson, L., & Switzer, D. (2012, March). *Challenges and benefits of implementing authentic inquiry-based in instruction through a research experience for teachers program*. Paper presented at the annual meeting of the National Association of Research in Science Teaching, Indianapolis, IN.
- Walters, K., Scheopner Torres, A., Smith, T., & Ford, J. (2014). *Gearing up to teach the Common Core State Standards for mathematics in the rural Northeast Region* (REL Report No. 2015-031). Washington, DC: Regional Educational Laboratory Northeast & Islands.
- Wheatley, K. F. (2002). The potential benefits of teacher efficacy doubts for educational reform. *Teaching and Teacher Education, 18*, 5–22.
- Wilson, S. M. (2013). Professional development for science teachers. *Science, 340*(6130), 310–313.
- Windschitl, M. (2002). Framing constructivism as the negotiation of dilemmas: An analysis of the conceptual, pedagogical, cultural, and political challenges facing teachers. *Review of Educational Research, 72*(2), 131–175.
- Windschitl, M. (2003). Inquiry projects in science teacher education: What can investigative experiences reveal about teacher thinking and eventual classroom practice. *Science Education, 87*(1), 12–143.
- Zwiers, J., O'Hara, S., & Pritchard, R. (2014). *Common Core standards in diverse classrooms: Essential practices for developing academic language and disciplinary literacy*. New York, NY: Stenhouse.

### Authors

SHERRY A. SOUTHERLAND is a professor in the School of Teacher Education and a codirector of FSU-Teach, G107 Stone Building, Florida State University, Tallahassee, FL 32306-4459; ssoutherland@fsu.edu. Her research focuses on the intersection of culture, emotions, and science learning, as well as the barriers and affordances to the reform of science teaching.

ELLEN M. GRANGER is the director of the Office of Science Teaching Activities and Outreach and a codirector of FSU-Teach, 1062 King Life Sciences, Florida State University, Tallahassee, FL 32306-4295; granger@bio.fsu.edu. Her research focuses on science teacher professional development.

ROXANNE HUGHES is the director for the Center for Integrating Research and Learning, National High Magnetic Field Laboratory, 1800 E. Paul Dirac Drive, Tallahassee, FL 32310-3706; hughes@magnet.fsu.edu. Her research focuses on the role of informal experiences in shaping learners' identity construction.

PATRICK ENDERLE is an assistant professor of science education, Department of Middle and Secondary Education, Georgia State University, 33 Gilmer Street SE, Atlanta, GA 30302-3965; penderle@gsu.edu. His research explores implementation of argumentation and other scientific practices in the science classroom and how to support science teachers in those efforts, as well as the development of science teacher professional development that encourages the enrichment of instructional beliefs and practices.

FENGFENG KE is a professor of the Educational Psychology and Learning Systems, 3205C Stone Building, Florida State University, Tallahassee, FL 32306-4453; fke@fsu.edu. Her current research focuses on game-based learning, immersive learning, computer supported collaborative learning, and inclusive design of e-learning.

KATRINA ROSELER is an assistant professor in the Division of Education, Chaminade University of Honolulu, 3140 Waiialae Avenue, Honolulu, HI 96816; katrina.roseler@chaminade.edu. Her research focuses on the influence of informal science experiences on the development of students' science identities and persistence in formal science studies.

YAVUZ SAKA is an associate professor of science education, Bulent Ecevit University, Eregli Egitim Fakultesi, Turkey; sakayavuz@gmail.com. His research focuses on enhancing the professional development of science teachers during the induction period.

MIRAY TEKKUMRU-KISA is an assistant professor of science education at the Florida Center for Research on STEM, G107 Stone Building, Florida State University, Tallahassee, FL 32306-4459; mtekkumrukisa@fsu.edu. Her research focuses on designing and studying the effectiveness of tools and environments (e.g., video cases, educative curriculum materials, professional development programs) for supporting science teachers' learning and instructional improvement.