Spring 4-15-2016

Building Exemplary Teaching Practices: Following the Paths of New Science Teachers

Elizabeth B. Lewis
University of Nebraska-Lincoln, elewis3@unl.edu

Ana Rivero
University of Nebraska-Lincoln

Aaron A. Musson
University of Nebraska-Lincoln, aaronmusson@gmail.com

Jia Lu
University of Nebraska-Lincoln

Lyrica Lucas
University of Nebraska-Lincoln, lyricalucas@huskers.unl.edu

Building Exemplary Teaching Practices: Following the Paths of New Science Teachers

Elizabeth Lewis, Ana Rivero, Aaron Musson, Jia Lu, and Lyrica Lucas
College of Education & Human Sciences
University of Nebraska-Lincoln
Lincoln, Nebraska

Abstract

There are few comprehensive studies of beginning science teachers that describe enacted teaching practices in terms of inquiry-based instruction, classroom discourse, assessment, and curricular choices, and explore how these factors interact with teaching self-efficacy. We conducted a 3-year, longitudinal study of four cohorts of master's level science teacher education program graduates. We coded and analyzed 319 science lessons of new teachers from student teaching to third year post-program to describe teachers' enacted practices and gathered annual teaching self-efficacy reports to examine teachers' beliefs. Our analysis resulted in key findings relevant to future programmatic improvements. First, when we reviewed specific inquiry-based teaching facets we found patterns indicating areas of growth and areas of challenge. Four areas of growth included teaching for knowledge acquisition, questioning level employed, conceptual development, and content depth. These aspects of teaching science were strongly addressed during the teacher education program. Some areas of challenge for these new science teachers included: using an inquiry-based order of instruction, promoting classroom interactions, accessing students' prior knowledge, and learner centrality in enacted curriculum. Second, we found that the number of years a teacher taught mattered when predicting overall self-efficacy, specifically for self-efficacy associated with student engagement and instructional strategies. Over time, it appears that the MAst teachers who have persisted through the induction period have maintained a positive outlook on their agency. We attribute the generally positive nature and stability of these beginning science teachers' self-efficacy to a rigorous teacher preparation program, but recognize that teachers could benefit from ongoing professional development in inquiry-based instruction, rich discourse strategies, and formative assessment.

Introduction: Beginning Science Teachers' Teaching Self-efficacy and Enacted Practices

There are few comprehensive studies of beginning science teachers (NRC, 2010) that describe enacted teaching practices in terms of inquiry-based instruction, classroom discourse, assessment, and curricular choices, and explore how these factors interact with teaching self-efficacy; our work addresses this gap. By understanding how individual aspects of teaching interact, we can better understand how to recruit teacher candidates and support them through their induction period to reduce attrition. Only 42% of middle school and 49% of high school science teachers have more than 10 years of teaching experience (Banilower, Smith, Weiss, Malzahn, Campbell, and Weis, 2013). Schools with higher percentages of students who qualify for free and reduced lunch are more likely than schools with fewer students in poverty to have less experienced teachers Banilower et al., 2013).

We define exemplary teaching as effective teaching practices. In a still-referenced vision of teacher preparation, Darling-Hammond & Bransford (2005) highlight three areas of skills, knowledge, and dispositions important for teachers: "(a) knowledge of **learners** and how they **learn and develop** within social contexts; (b) conceptions of **curriculum content and goals**: an understanding of the subject matter and skills to be taught in light of the social purpose of education; and (c) an understanding of **teaching** in light of the content and learners to be taught, as informed by assessment and supported by classroom environments" (p. 11, Figure 1). Bianchini (2012) found that little is known about the science teaching induction period, recommending more studies that: (a) follow beginning science teachers from preservice teacher education into classroom practice and (b) trace connections, or lack of, across induction training, beginning teachers' classroom practices, and student learning. Our research contributes to understanding how to construct effective science teacher education programs (TEPs) that result in teachers who can address national science education standards to educate youth to be scientifically literate citizens, as well as encourage more students to pursue STEM careers to meet the national call for a more highly qualified STEM workforce. With new national science education standards (NGSS Lead States, 2013), it is critical we understand how to educate science teachers capable of advancing these priorities.

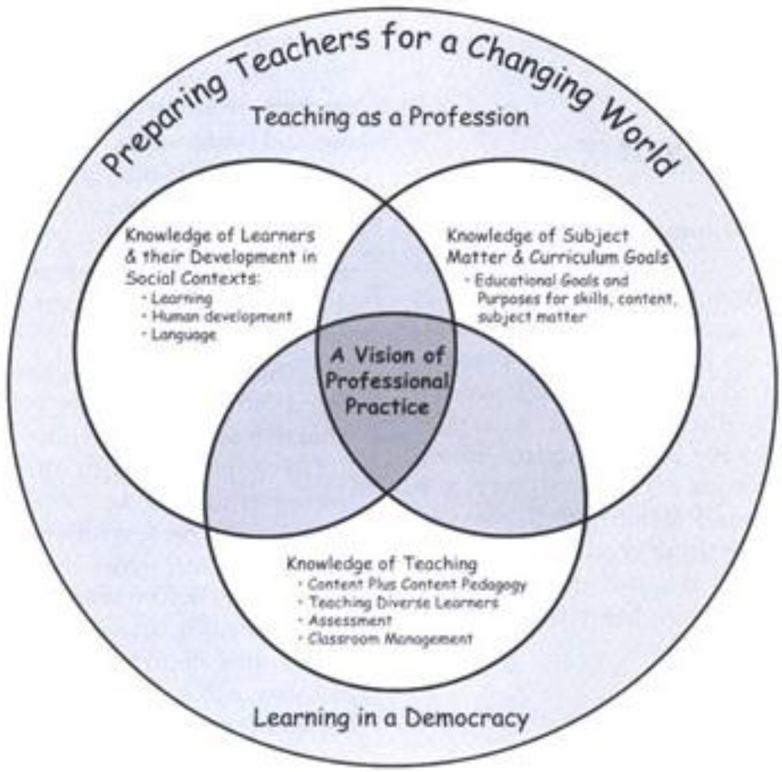

Figure 1. Vision of professional practice for teachers (taken from Darling-Hammond & Bransford, 2005, p. 11).

Background Literature

We briefly summarize some of the research and theories that have guided our work. In the interest of space we have only included a few examples of foundational work in these areas and/or recent work in science education. All of the literature review for all five "papers" are included in this section in the same order as the results are presented.

Teaching Self-efficacy. Pajares (1992) argued research agendas must attend to PSTs' beliefs as a means for informing educational practice. All PSTs' learning is filtered through their beliefs and perspectives, which shape their teaching philosophy and instructional practices. Bryan (2012) noted the large amount of literature "that establishes that teachers are creative, intelligent decision makers who hold complex systems of beliefs that influence how they view students, themselves, and science" (p. 477-478). Teachers' beliefs have been studied for many decades and beginning with Bandura's (1997) work others have been interested in learning more about how teachers' sense of self and their teaching self-efficacy may affect their curricular and instructional choices (Jones & Leagon, 2014; Tschannon-Moran & Hoy, 2001). High levels of teaching self-efficacy has been shown to be an indicator of more innovative teaching (Guskey, 1988) and to contribute to higher student achievement (Evans, 2011). Teaching self-efficacy is important in science education because teachers must be equipped to problem-solve student learning and fundamentally believe that what they are doing will help their students learn better. Those teachers who understand how students learn and have high teaching self-efficacy will have a better chance of helping students become scientifically literate because they will not rest until they have done everything they can to problem-solve student learning. However, it is important to note that sometimes teachers have conflicting, or competing belief sets (Crawford, 2007), such as school culture (McGinnis, Parker, and Graeber, 2004), that can disrupt even a positive self-efficacy for enacting inquiry-based science instruction.

Science Teachers' Instruction. Beginning science teachers need to expand their abilities to develop and implement inquiry-based lessons. This is one of the aspects of learning to teach science that have been the focus of the current research literature (NGSS Lead States, 2013) that guide policy in the United States. The inquiry approach to teaching and learning is promoted in science teacher preparation programs in response to science education research literature and recommendations drafted in various versions of standards for teaching science (NGSS Lead States, 2013; NRC, 2010; NRC, 1996). Supovitz, Mayer, and Kahle (2000) defined inquiry-based instruction as "a student-centered pedagogy that uses purposeful, extended investigations set in the context of real-life problems as both a means for increasing student capacities and as a feedback loop for increasing teachers' insights into student thought processes" (pp. 331-356).

Since science teachers employ a variety of instructional methods and strategies, an examination of student-centered instructional practices can serve as a window to understanding the quality of inquiry-based instruction.

General instructional methods and strategies used in science classes can be viewed in terms of the amount of direct control that teachers and instructors have over their implementation (Treagust & Tsui, 2014; Treagust, 2010). In learning environments guided by the inquiry approach, instructional practices characteristically depart from traditional teacher-centered methods. Teachers are more likely to deliberately design and select learner-centered methods and strategies that encourage explorations and questioning. This proclivity to devote more time on student learning is a quality of efficacious teachers (Woolfolk & Margetts, 2007). Thus, teacher self-efficacy indicators (Tschannen-Moran & Woolfolk-Hoy, 2001) and instructional factors (Marshall, Smart, & Horton, 2010) may converge and influence enacted practices in science classrooms.

Discourse in the Science Classroom. Since the publication of the National Science Education Standards (NRC, 1996) and the Benchmarks for Scientific Literacy (AAAS, 1993), "inquiry" has remained a central term in science education in the United States. In *Inquiry and the National Science Education Standards* (NRC, 2000), essential features of classroom inquiry are described through what the learners are doing. The more the learners are engaged in scientifically-oriented questions and/or communicating their scientific understanding, the more inquiry-oriented the class is likely to be. Viewing language and communication as essential elements in science learning is echoed in the emphasis on language intensive disciplinary practices across new standards such as the *Common Core* and *Next Generation Science Standards* (NGSS) (Lee, 2013). Viewing language as a system of resources for meaning-making and communication as a social process is a change that came with Lemke's (1990) seminal study on the limited ways science was talked about in secondary science classrooms. Lemke proposed that science education should enable students to become "fluent speaker of science" and "we have to learn to see science teaching as a social process and to bring students...into this community of people who talk science" (p. x). By "talking science" he meant not just talking about science, but also "doing science through the medium of language" (p. ix). Lemke's reconceptualization of science learning as doing science through language in a community of speakers of science places language at the center of science learning. Such reconceptualization challenges the traditional use of language as a tool for transmission of information about natural phenomena; it leads us to rethink language as "an interpretive system" (Sutton, 1996) where meaning-making, exploring and persuading happen. In other words, learning science is developing a repertoire of discursive practices to engage in scientific knowledge and practices (Kelly, 2008). Lemke's reconceptualization of science learning also incorporates a sociocultural view on meaning making that redefines the role of teachers. Viewing meaning as constructed among people through dialogical process, Mortimer and Scott (2003) describe the teacher's role as a mediator who introduces, frames, shapes, and evaluates dialogues about natural phenomena to develop a rich environment for students to engage with scientific ideas and internalize knowledge constructed by teachers and students in this process.

Research on classroom discourse in science learning has identified questioning as a common pedagogical practice to facilitate science learning. Metacognitive questions that call the learners' attention to their own thinking and their own knowledge are used to both assist and assess student learning. When teachers ask students metacognitive questions, they are able to understand what students understand (Duckworth, as cited in Cazden, 2001). In inquiry-based science classrooms, authentic questions (Roth 1996; Marshall et al., 2008), or questions without a preconceived response are used to create a student-centered learning environment as opposed to the traditional Initiate-Response-Evaluation whole group discussion model (Lemke, 1990).

Language and communication in science classroom is a key aspect of equity and underrepresented students' access to scientific knowledge. Since science classroom discourse practices are often based on taken-for-granted assumption about ways of talking science, classroom discourse practices can serve to build knowledge and affiliation or limit participation and access depending on students' previous experiences, cultural assumptions, and worldviews (Lee, 1999). Studies of classroom interaction also show that compared with their male peers, female students have fewer opportunities to interact with the teacher, to be challenged by more complex questions, and to practice paradigmatic discourse (Kelly, 2008).

Assessment Practices. Assessment has been an essential element in education inside the classroom. The term *assessment* refers to all the activities that provide information about students' learning. This information is useful for both students and teachers. For students, it is a way to measure their own development, strengths, and limitations to increasing their learning. For teachers, it provides feedback to inform their teaching and choice of learning activities, curriculum, and instruction to support their students (Black and Williams, 1998). Wiggins (1998) described the two main functions of educative assessment as: (1) to teach (i.e., to be part of the instructional activities); and (2) to provide feedback about students' performance.

Since the beginning of the 20th century, Dewey and other progressive educators considered schools as places to develop students' thinking using inquiry methods (Dewey, 1910). In science education, several efforts have been conducted to increase inquiry-teaching. For example, the *Next Generation Science Standards* (NGSS) include scientific and engineering practices to engage students in the world of scientific activities, elicit their reasoning, and help them to apply scientific principles, (Haag & Megowan, 2015; Osborne, 2014). These practices are a fundamental part of the science curriculum and students' expected performances. Assessment practices should be tightly linked to curriculum and instruction (Osborne, 2007; Wilson & Bertenthal, 2005). Therefore, assessment should focus on those scientific practices and contribute to their development.

Consequently, there is a call for science education for assessment practices to promote inquiry and scientific reasoning (Blanchard, Southerland, Osborne, Sampson, Annetta, & Granger, 2010; Pellegrino, 2012; Wiliam, 2007; Black & Wiliam, 1998; Marshall & Drummond, 2006; Songer & Ruiz Primo, 2012). An inquiry-oriented assessment integrates assessment with the instruction of scientific practices to develop students' learning and inform teaching. It is known as assessment *for* learning or formative assessment (Bell & Cowie, 2001; Marshall & Drummond, 2006; Marzano 2010). Assessment for learning in an inquiry lesson should transform students into independent learners (Marshall & Drummond, 2006). It should provide tools to transfer scientific knowledge and skills into their lives. It should focus on student thinking (Coffey, Hammer, Levin, & Grant, 2011), which is not easy to teach (Furtak, Morrison, & Krogg, 2014). Science teachers must be active, creative, responsive to students' needs, reflective of their teaching practices and have flexibility to reorganize their curriculum while teaching (Furtak, Morrison, & Kroog, 2014). They require sufficient content knowledge and pedagogical knowledge (Nilsson, 2013). Bell and Cowie (2001) suggest that effective assessment for learning is especially difficult for novice teachers. This study seeks to contribute to our understanding of assessment for learning in beginning science teachers' lessons.

Teachers' Curricular Choices. According to Linn, Songer and Eylon (1996) there have been three historical periods that have reflected the degree of collaboration among groups concerned with science education, but it was not until 1975, the start of the so-called "partnership period" that collaboration among experts began to occur (DeBoer, 2014, p.573). However, DoBoer (2014) comments that since the beginning of the partnership period research on curriculum has increased, but that even now most curriculum materials are not research-based. For example, only about 3% of high school classrooms use materials that have been supported by NSF funding, which require a strong theoretical foundation for learning (Banilower, et al., 2013).

Teachers' choice of curriculum controls students' opportunities to learn science. The depth of the science content varies from lesson to lesson, but should be sufficiently rigorous to challenge students at the cognitive level that they current occupy. How students interact with science lessons and activities and the degree to which inquiry-based curriculum is provided that involves students' executive control of their learning have been shown to be important cognitive aspects of learning new concepts and ideas (NRC, 2005). Additionally, teachers' choice of curriculum that connects to socioscientific issues can be more engaging to students and promote scientific literacy (Zeidler, 2014).

The 2012 National Survey of Science and Mathematics Education (NSSME), funded by the National Science Foundation, revealed that at least once a week 49% of middle school and 65% of high school novice science teachers regularly had their students engage in hands-on/laboratory activities, and at least once a week 53% of middle school teachers and 35% of high school teachers required students to read from their science text aloud or to themselves, and occasionally engaged students in project-based learning (MS=20%; HS=20%) (Banilower, Trygstad, and Smith, 2015).

Research Methods

For this investigation we adopted an exploratory, multi-method approach to investigating beginning science teachers' teaching self-efficacy and enacted practices. We used a validated survey, adapted interview protocols, and engaged in regular classroom observations with a validated instrument to code inquiry-based science instruction. The context of and participants in our studies that led to this NARST paper set are described below.

Context. We conducted a three academic years (2012-2015), longitudinal study of secondary science teacher program graduates from a large Midwestern (U.S.) 4-year state university. The program only recruited teacher candidates who had earned at least a bachelor's degree in a scientific field, thus meeting the federal definition of a "highly-qualified" teacher. The program culminated in both initial secondary science certification (Table 1) and a 42-credit hour master of arts in science teaching. The Master of Arts in teaching (MAT) program is a 14-month, 42-credit hour program that provides a pathway for recent science graduates and practicing scientists to obtain secondary science certification. The program incorporates three major threads: coursework required for teacher certification, supporting graduate-level courses that include a capstone action research project, and extensive (600+ hours) clinical experiences. MAT students begin as a cohort in May and graduate in August of the following year (the specific teacher education program details and teachers' content knowledge were presented in a previous NARST conference presentation, Lewis, Musson, and Lu, 2014).

Once successful teachers left the MAT program they were certified and began their new teaching positions. Many teachers took teaching positions in high-needs school districts as they were required by the NSF Noyce stipend they received to do the MAT program to teach for two years in such a district. The MAT program and field placement coordinators made every effort to place preservice teachers in practicum and student teaching situations in diverse schools to try to prepare them for teaching students with a wide range of learning needs, including English learners.

Table 1. *Science Teaching Endorsements of Teacher Graduates.*

Cohort	Median age range (years)	Average time between degrees (years)	Single-subject Endorsements (Required minimum: 24 credit hours)			
			Biology	**Chemistry**	**Earth**	**Physics**
MAT-1 (*n*=14)	27.8 (22-46)	5.3	8	4	0	2
MAT-2 (*n*=16)	24.3 (22-53)	3.0	15	6	1	0
MAT-3 (*n*=11)	26.6 (22-42)	2.6	7	4	1	0
MAT-4 (*n*=10)	23.1 (23-43)	2.8	6	4	1	2
		*** Total:**	**36**	**18**	**3**	**4**

Note: * *Individual teachers may have earned more than 1 single-subject science teaching endorsement.*

Teaching Self-efficacy Data. We evaluated the teacher education program graduates at the end of their student teaching (ST, *n*= 41), and each year thereafter (Y1, *n* = 24; Y2, *n* =20; Y3, *n*= 8). We used the *Teacher Sense of Efficacy Scale* (TSES), a 24-item survey instrument with a five-point scale developed by Tschannon-Moran and Hoy (2001), to investigate teacher efficacy in three areas: (a) Student Engagement (SE), (b) Classroom Management (CM), and (c) Instructional Strategies (IS). We examined the teachers' changing self-efficacy using a multivariate analysis of variance (MANOVA). Our three outcome variables were the three subscales on the instrument and we used number of years of teaching experience to predict change across the multiple outcome measures.

Classroom Data. We analyzed 319 science lessons of induction phase teachers from their student teaching placements to their third year after completing the MAT education program (four cohorts from 2012 to 2015). Our dataset included 71 lessons by student teachers (*n*=33), 116 by first- (*n*=26), 95 by second- (*n*=19), and 37 by third-year teachers (*n*=6). We regularly visited teachers' classrooms by arranging visits that did not include times when teachers were administering tests or spending the whole class period showing a video. Teachers approved our visits and we tried to visit to see a range of different lessons, if possible. Five researchers observed and coded lessons using the EQUIP instrument (Marshall, Horton, Smart, & Llewellyn, 2008) to measure the quality of inquiry-based instruction in middle and high school science classrooms. Regular calibration of the research team occurred throughout the three academic years of data collection by using videos of lesson to learn to use the instrument and periodically conducting observations in pairs (with all possible combinations of observers) to ensure that the team's calibration had not drifted. The EQUIP instrument employs a scale of 1 to 4 to describe the degree of inquiry in a lesson. Level 1 corresponds to "pre-inquiry" (i.e., a teacher-centered lesson) and 4 to "exemplary inquiry" (i.e., an open-ended and engaging student-centered lesson). We used frequency counts in "proficient" and "exemplary" inquiry" (Levels 3 and 4) to analyze the five constructs measured with the EQUIP assessment factors.

Research Questions. We explored the teaching practices of beginning science teachers' practices with respect to exemplary, reform-based instruction using the following research questions:

1. How does the teaching self-efficacy (specifically, in terms of student engagement, classroom management, and instructional strategies) of beginning science teachers change over time, if at all? (Paper 1)
2. What, if any, changes in inquiry-based teaching practices (specifically, instruction, discourse, assessment, and curriculum) have occurred over time as induction-level (Years 1-3) science teachers gained experience? (Papers 1-5)

Paper #1: Overall Patterns of Science Teaching Self-efficacy and Teaching Practices

Results

Teaching Self-Efficacy. We examined the teachers' teaching self-efficacy using a multivariate analysis of variance (MANOVA). Our three outcome variables were the three subscales on the instrument, self-efficacy regarding: (a) student engagement, (b) instructional strategies, and (c) classroom management (Table 2). We used number of years of teaching experience to predict change across the multiple outcome measures. We discovered that time spent teaching accounted for average differences across the three measures, Wilk's Lambda (9, 211) = 2.02, p=0.04. In simple follow-up tests using a Bonferonni adjustment, we found that there were only significant changes in student engagement ($F (3, 89) = 4.54$, $p < 0.01$) and instructional strategies ($F (3, 89) = 3.17$, $p = 0.03$) (not classroom management ($F (3, 89) = 1.18$, $p = 0.32$). Going further, we isolated the pairwise comparisons for number of years teaching (0-3) with the two subscales, student engagement and instructional strategies, for which there were statistically significant results. We again adjusted our *p*-values for multiplicity issues, and found statistically significant differences between student teaching and Years 1 and 2 of teaching for student engagement, and only for the difference between student teaching and Year 1 of teaching for instructional strategies. Of note is that no pairwise comparisons that included teachers with three years of teaching yielded significant results.

To summarize, our findings were that the number of years a teacher taught mattered when predicting overall self-efficacy, and specifically for self-efficacy associated with student engagement and instructional strategies. Longitudinal comparisons were only meaningful when we used the teachers as their own controls (i.e., their responses at end of their student teaching). This suggested that the measurement instrument was not sensitive to changes in teachers' self-efficacy after two or more years of having exited the MAT program. Over time, it appears as if the MAT teachers who have persisted through the induction period have maintained a generally positive outlook on their own agency (i.e., they can do "some" to "quite a bit" to affect positive change) in these three areas of teaching, remaining generally optimistic even after their first year of teaching. Overall, these beginning science teachers appear to think that their actions can result in increasing student engagement, keeping classrooms running smoothly, and implementing effective instructional strategies. We attribute the positive nature and stability of these beginning science teachers' self-efficacy to intellectually strong teacher candidates and a rigorous teacher preparation program. In our previous work we have been able to show that the number of credit hours a teacher has in their area of certification (data for chemistry and physics only) predicts a higher score on tests of misconceptions, i.e., the teacher holds fewer misconceptions (Lewis et al, 2014). These data will be incorporated with the full set of data in our next study to build a structural equation model (see conclusions section at end of paper set for future work).

Table 2. *Average Teaching Self-efficacy of Teacher Graduates.*

	Post-Student Teaching	**Post-Year 1**	**Post-Year 2**	**Post-Year 3**
Number of teachers	*41*	*24*	*20*	*8*
Student Engagement Mean	3.84	3.54 *	3.49 *	3.56
SD	0.46	0.36	0.35	0.39
Classroom Management Mean	4.05	3.76	3.84	3.97
SD	0.42	0.37	0.34	0.39
Instructional Strategies Mean	4.15	3.94 *	4.01	3.92
SD	0.49	0.47	0.51	0.50

Note: * = statistically significant difference.

Summary of all science lessons. We generated a sample of 319 observations of science lessons during multiple years of teaching by beginning science teachers (Table 3). We used the EQUIP instrument to code these observations of teachers from student teaching to teachers' third year teaching. Overall, the areas that appeared to show the greatest growth toward inquiry-based instruction as teachers gained more experience were on the instructional factors and discourse factors scales. Some more modest growth was observed on the curriculum factors scale.

When we reviewed specific items on the EQUIP there is a clearer pattern of growth and areas of challenge. We have selected representative items to illustrate this, but in each of the other papers in this set we focus on the individual item score results from our observations as a way to better understand specific curricular aspects of these enacted lessons. Examples of four areas of steady growth toward more inquiry-based practices included: (a) teaching for knowledge acquisition, (b) questioning level employed, (c) conceptual development, and (d) content depth. These areas were strongly addressed during the MAT program. Some areas of challenge included: (a) order of inquiry-based instruction, (b) classroom interactions, (c) accessing students' prior knowledge, and (d) learner centrality in selected curriculum (Figure 2). While the 5E model of inquiry-based instruction was used to frame science teaching methods courses in the MAT program, ongoing professional development may be needed to support further growth in these beginning science teachers. The most persistently lowest scoring aspect, assessment, showed little growth from first to third year teaching. This suggested that a better effort may be needed, on our part, to document these teachers' practices of assessment and/or teachers need more professional development to encourage the use of more standards-aligned, formative and summative assessment practices. When we discuss professional development needs with graduates of this program they often mention their desire to know more about effective assessment practices.

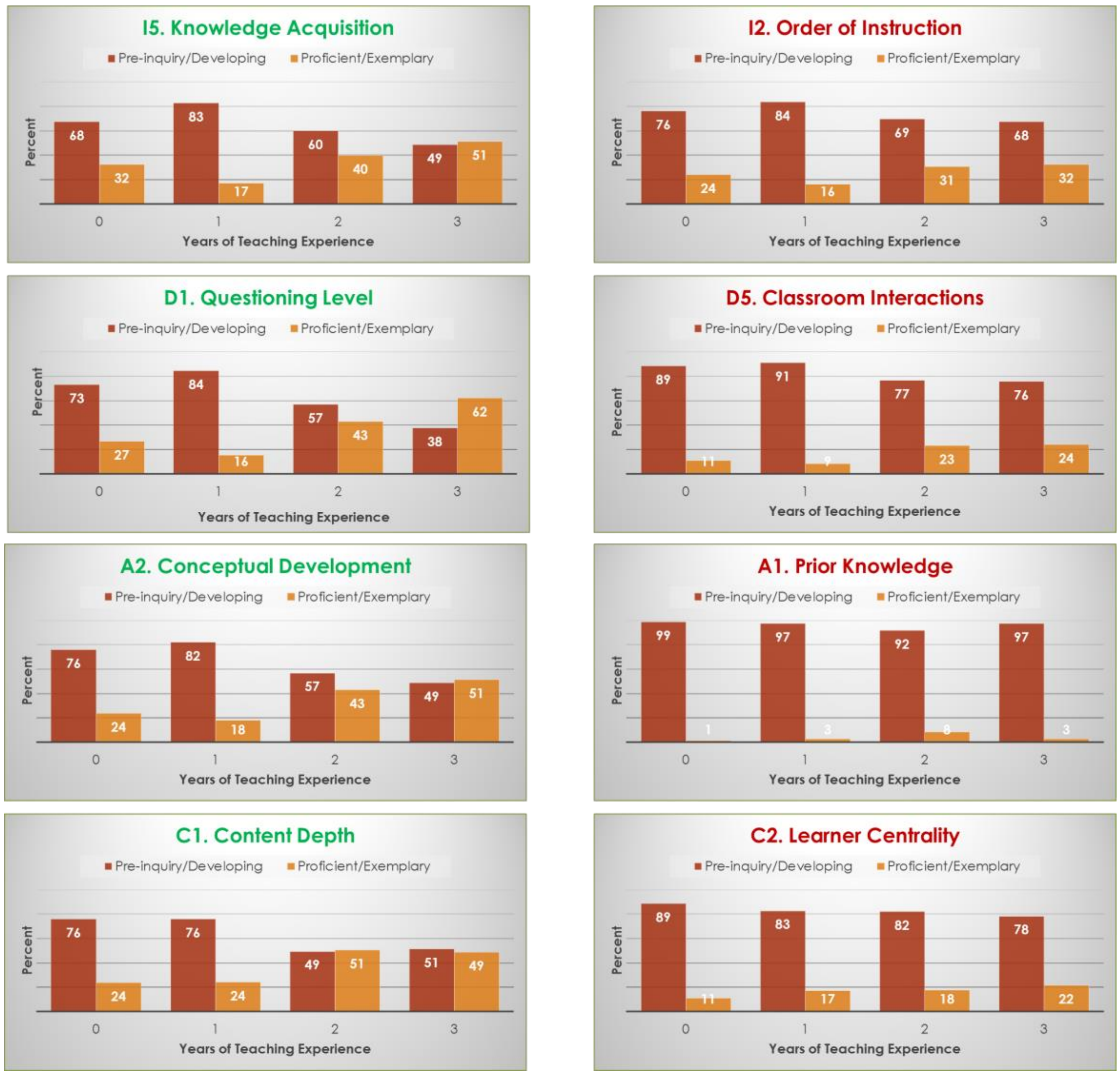

Figure 2. EQUIP-identified areas of most growth (graphs on the left) toward inquiry-based teaching practices and areas in most need of professional development (graphs on the right).

Table 3. *Descriptive Statistics of Teachers' Average Enacted Curricular Practices Measured with EQUIP.*

Teaching Phase	Student Teaching		Induction Year 1				Induction Year 2				Induction Year 3			
Time point	Sem 0 (Spring)		Sem 1 (Fall)		Sem 2 (Spring)		Sem 3 (Fall)		Sem 4 (Spring)		Sem 5 (Fall)		Sem 6 (Spring)	
# of Lessons:	71		38		78		39		56		14		23	
# of Teachers:	33		22		26		18		19		6		6	
EQUIP Scale actors	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Instructional	2.20	0.77	2.06	0.75	2.07	0.81	2.31	0.95	2.21	0.84	2.31	0.93	2.59	0.87
Discourse	2.04	0.65	1.95	0.61	1.87	0.69	2.20	0.72	2.22	0.70	2.19	0.77	2.56	0.59
Assessment	1.73	0.67	1.72	0.62	1.74	0.66	2.00	0.78	1.90	0.77	1.87	0.72	2.03	0.78
Curriculum	1.92	0.72	1.97	0.68	2.02	0.73	2.13	0.81	2.08	0.80	2.17	0.80	2.21	0.86
EQUIP Total:	1.98	0.72	1.92	0.68	1.92	0.74	2.16	0.83	2.10	0.79	2.13	0.82	2.35	0.81

Paper #2: Instructional Factors and Teaching Self-efficacy of New Science Teachers

Results

From the EQUIP data we identified trends in five constructs: (a) instructional strategies, (b) order of instruction, (c) teacher role, (d) student role, and (d) knowledge acquisition. These items compose one EQUIP scale, *instructional factors*. The 4-point EQUIP scale measures the level of inquiry instruction enacted by a teacher from pre-inquiry (Level 1) to exemplary inquiry (Level 4). For instance, in terms of instructional strategies, a teacher may be observed to predominantly lecture to cover content (Level 1) or occasionally lecture but use classroom activities that promoted strong conceptual understanding (Level 4).

In our analysis of 319 observed science lessons, we found that Year 3 teachers applied more proficient or exemplary inquiry-based approaches when compared with preservice, Year 1, and Year 2 teachers (Table 4). The level of inquiry-based instruction declined slightly in the first year of teaching relative to preservice practice, and appeared to increase and trend toward more learner-centered methods thereafter. The discrepancy in the application of inquiry-based practices between preservice and Year 1 teachers may be explained by the steady and easier access to various resources embedded in the teacher preparation program through taking two science teaching methods courses and a student teaching seminar in succession, as well as being supported by an experienced cooperating teacher. Preservice teachers in the program were explicitly encouraged to design lessons following the 5E (i.e., *Engage, Explore, Explain, Elaborate, and Evaluate*) model that subscribes to "science as inquiry" thinking. It appears that losing these supports afforded by the teacher preparation program during their first year of teaching may have impacted the quality of inquiry-based instruction. Among these five instructional factors, *order of instruction* emerged as an area in need of continued emphasis. On the other hand, teachers showed a steady growth from Year 1 to Year 3 in all of the other four constructs of instructional factors in the EQUIP scale.

To guide our inquiry on how teacher self-efficacy and instructional factors may converge and influence enacted practices in science classrooms, we created a matrix table (Table 5) of the constructs of instructional factors and self-efficacy items from the EQUIP scale and the TSES, respectively. The TSES has the three subscales on self-efficacy regarding: (a) student engagement, (b) instructional strategies, and (c) classroom management. In Paper #1 of this set (also presented in Lewis, Musson, Rivero, Lu, and Lucas, 2015), we found that the number of years a teacher taught mattered when predicting overall self-efficacy, and specifically for self-efficacy associated with student engagement and instructional strategies. In our matrix comparing instructional factors and self-efficacy reports, we selected items from the instructional strategies and student engagement of the TSES. Table 5 shows that the substantial change in self-efficacy occurs in Year 2. This is supported by our previous finding that the measurement instrument was not sensitive to changes in teachers' self-efficacy after 2 or more years of having exited our teacher education program (Lewis et al., 2015). Teachers appeared to rate and label their efficacy as consistently high from Year 2 to Year 3.

Table 4. *Effective Aspects of Instruction: Percentage of Observed Science Lessons at "Proficient" or "Exemplary" Levels of Inquiry* (n=319 lessons)*

Instructional Factors	**Student Teaching** % (n=71)	**Year 1 %** (*n*=116)	**Year 2 %** (*n*=95)	**Year 3 %** (*n*=37)	**Mean %** (with student teaching)	**Induction (Years 1-3)** Mean %
I1: Instructional Strategies	41	32	45	49	42	42
I2: Order of Instruction	24	16	31	32	26	26
I3: Teacher Role	34	32	39	51	39	41
I4: Student Role	44	37	47	59	47	48
I5: Knowledge Acquisition	32	17	40	51	35	36

* Note: *Proficient scored a "3" and "Exemplary" scored a "4" on the EQUIP instrument.*

Table 5. *Percentage of Teachers Reporting High Levels of Efficacy (4 = Quite a bit and 5 = A great deal)*

EQUIP Item: Instructional Factors	**TSES Item: Teaching Self-efficacy**	**Year 1 %** (*n*=23)	**Year 2 %** (*n*=21)	**Year 3 %** (*n*=9)
I1: Instructional Strategies	Q10: How much can you gauge student comprehension of what you have taught?	70	81	100
	Q17: How much can you do to adjust your lessons to the proper level for individual students?	52	48	44
	Q23: How well can you implement alternative strategies in your classroom?	35	57	67
	Q24: How well can you provide appropriate challenges for very capable students?	35	52	56
I2: Order of Instruction	Q7: How well can you respond to difficult questions from your students?	100	95	100
	Q11: To what extent can you craft good questions for your students?	65	76	89

	Q20: To what extent can you provide an alternative explanation or example when students are confused?	78	95	100
I3: Teacher Role	Q17: How much can you do to adjust your lessons to the proper level for individual students?	52	48	44
	Q23: How well can you implement alternative strategies in your classroom?	35	57	67
	Q24: How well can you provide appropriate challenges for very capable students?	35	52	56
I4: Student Role	Q17: How much can you do to adjust your lessons to the proper level for individual students?	52	48	44
	Q24: How well can you provide appropriate challenges for very capable students?	35	52	56
I5: Knowledge Acquisition	Q2: How much can you do to help your students think critically?	74	76	78
	Q10: How much can you gauge student comprehension of what you have taught?	70	81	100
	Q12: How much can you do to foster student creativity?	39	43	67
	Q18: How much can you use a variety of teaching strategies?	65	52	78

Discussion

In general, the increasing enactment of inquiry-based practices along with years of teaching experience coincides with improvements of some aspects of teachers' self-efficacy. Our findings support claims that field experiences helps teachers to develop more sophisticated ideas about science instruction and acquire self-efficacy as science teachers (Davis, Petish, & Smithey, 2006). Increased used of inquiry-based instruction (i.e., in terms of the constructs in the *instructional factors* scale) among in-service teachers with longer field experience appear to be concurrent with increasing self-efficacy in some aspects such as gauging student comprehension, implementing alternative strategies, providing appropriate challenges for very capable students, crafting good questions, providing an alternative explanation or example, fostering student creativity, and using a variety of teaching strategies.

We find it curious and telling that although teachers report high self-efficacy in responding to students' questions, crafting good questions, and providing alternative explanations and examples, beginning science teachers appeared to be predisposed to explain concepts and provide limited opportunities for students to explore and arrive at their own conceptual explanations based on our class observations and therefore score lower in the measure for *order of instruction* in the EQUIP scale. These findings remind us that teachers' sense of their own self-efficacy is not a purely objective, or independent, measure of actual competence in practicing inquiry-based instruction. Therefore, comparing EQUIP scores generated from classroom observations along with teachers' report of efficacy allows us to probe areas where teachers' self-evaluation converge with the findings from our classroom observations. Other aspect of teaching and beliefs, such as teachers' perceptions of school policy and culture may also affect teachers' instructional decisions.

While it is important to note the areas of instruction where teachers believe they are doing well based on their self-efficacy assessment, but score low in the corresponding construct in the EQUIP scale, items in the TSES that teachers rated as low are also revealing. In Table 4, we observe that although inquiry-based practices are most likely increasing over time, this change is gradual and teachers do not appear to demonstrate exemplary inquiry very often until they reach Year 3. The self-efficacy reports show that Year 1 and Year 2 science teachers reported lower levels of self-efficacy in adjusting lessons to the proper level for individual students despite gaining more experience. This may be due to the fact that teachers' teaching assignments change from year to year, or even if teachers are teaching the same courses they sometimes report that while they are less stressed they are still figuring out what types of instruction work best with their students. Our teachers have reported both of these situations to us when we visit their classrooms, but we have yet to make a formal study of the degree to which these issues affect their instruction.

Our findings about teachers' instruction point to a clear need for support during the early years of teaching and for beginning teachers to gain a better understanding of how the *order of instruction* affects students' access to and the quality of inquiry in the classroom. Furthermore, understanding the factors influencing the development of teaching self-efficacy through research is necessary to support sources of teacher learning and growth. Although this study has identified the areas of challenge in terms of inquiry-based instruction and revealed that first-year teachers manifest a tendency to adopt a traditional teacher-centered approach to teaching, we have not yet studied how teachers' involvement in professional development could address these areas of need. While the 5E model of inquiry-based instruction was used to frame science teaching methods courses in the graduate program taken by participating science teachers, our findings showed that ongoing professional development would be needed to support further growth in these beginning science teachers.

Paper #3: Discourse in Beginning Science Teachers' Classrooms

Conceptual Framework

We approach acts of teaching and learning through the sociocultural model described by Mortimer and Scott (2003). Within this model the science teacher acts as a mediator, and each learning event happens in three stages: the teacher (1) makes ideas "available on the social plane of the classroom," (2) monitors and assists students as they rehearse and internalize the ideas, and (3) helps students apply the scientific ideas beyond the lesson (Mortimer and Scott, 2003, p. 17). Mortimer and Scott's description of a learning event is congruent with the 5E teaching model and the EQUIP instrument describes to what extent teachers apply inquiry practices, based largely on the 5E model (Marshall, Horton, Smart, & Llewellyn, 2008). The *Engage* and *Explore* components of the 5E model may be seen as making "the scientific ideas available" (Mortimer and Scott, 2003, p. 17) as students "engage with a new concept [and] make connections between past and present learning experiences" (Bybee, et al, 2006). Likewise, students rehearse and demonstrate their understanding with their teacher, and construct working *Explanations* and propose possible *Elaboration* or applications of the phenomenon or concept as they develop their "scientific story" (Mortimer and Scott, 2003, p. 18). The teacher's role as a mediator reflects the *Evaluation* component of the 5E model, as the teacher monitors and supports students' efforts to construct meaning from the learning event.

Paper-specific Research Questions

In our paper we explore the following two sub-questions:

1. In what ways, and to what extent, do science teachers' discourse practices change during their induction phase?
2. How do science teachers mediate science learning through classroom discourse, especially through questioning?

Methods

To answer the first question, we analyzed a total of 319 observed science lessons using the EQUIP instrument (Marshall et al., 2008). The discourse factors scale on the EQUIP was used to describe the science classroom discourse and to document teachers' changes, if any, in discourse practices as they gained more experience. There are five constructs on the discourse factors scale with a scale of 1 to 4 to describe the degree of inquiry in relation to classroom discourse. Of the five constructs, the first three are concerned with teacher questioning and the fourth and fifth are descriptors of the dynamics of communication in science classroom. Data were also collected through a post-year belief survey (Tschannan-Moran & Hoy, 2001) and a self-developed class activity rubric was also analyzed for triangulation.

To answer the second question, a member of our research team, Aaron Musson, arranged to video record representative lessons of two third-year physical science teachers. We selected Carl and Kari based on their similar content-area preparation (Kari holds a master's degree in chemistry, and Carl completed about half the coursework for a master's degree in astronomy), their status as career changers, but their distinctly different teaching environments. Aaron observed and recorded 12 lessons taught by each participant and coordinated with Carl and Kari to observe a purposeful sample of a variety of lessons that included labs, demonstrations, and lectures. He interviewed Carl and Kari after each observed lesson to explain their decisions about their choice of questions, the resulting student-teacher dialogue, and class discourse "in the moment." For each participant, we selected video clips of six different interactions, and asked our participants to "talk us through" their decisions. Aaron interviewed Carl and Kari using a version of the teaching beliefs interview protocol developed by Luft and Rohrig (2007). Additionally, we recorded and analyzed our participants' statements as they watched video recordings of their lessons. Finally, we extracted and analyzed the teachers' explicit statements related to their beliefs about teaching and learning.

Results: Changes in Classroom Discourse Practices

Preliminary analysis of the entire set of 319 science lessons shows an increase in all five constructs of discourse factors (See Table 6), which we interpret to mean that the teachers' classroom discourse has become more inquiry-based and student-centered. There exists a noticeable increase in the three constructs (D1-D3) related to questioning, with about 15% of the lessons demonstrating higher-level inquiry during the first year as opposed to around 50% of the lessons during the third year. With regard to the communication dynamics, the increase appears to be a slow, but steady increase on both constructs (D4 & D5). Out of all five discourse constructs measured using the EQUIP scale, classroom interactions (D5) is the area in which the least amount of change has occurred across the years, from 9% in Year 1 to 26% in Year 3. In general, our results indicate that classrooms of induction teachers become more student-centered and more inquiry-based as these teachers became more experienced. This trend is supported by some data form the teachers' self-efficacy survey. Table 7 is a summary of science teachers' response to Question 11 on the survey: *"To what extent can you craft good questions for your students?"* As shown in the table, the percentage of teachers who answered "some" decreased from 36% in Year 1 to 13% in Year 3 while those who chose "quite a bit" or "a great deal" increased from 63% to 88%.

Table 6. *Percentage of Observed Science Lessons Reaching Proficient or Exemplary Levels of Inquiry (n=319 lessons, not teachers)*

Discourse Factors	**Student Teaching % (*n*=71)**	**Year 1 %** (*n*=116)	**Year 2 %** (*n*=95)	**Year 3 %** (*n*=37)	**Mean %** (with student teaching)	**Induction (Years 1-3) Mean %**
D1: Questioning Level	27	15	44	61	37	40
D2: Complexity of Questions	24	15	37	45	30	32
D3: Questioning Ecology	24	15	40	47	33	34
D4: Communication Pattern	20	16	29	39	26	28
D5: Classroom Interactions	11	9	23	26	17	19

* *Note: Proficient scored a "3" and "Exemplary" scored a "4" on the EQUIP instrument.*

Table 7. Teacher responses to Question 11: *"To what extent can you craft good questions for your students?"* (*n= teachers who completed survey at the end of teaching year*)

Self-efficacy Q11	**% Nothing**	**% Very little**	**% Some**	**% Quite a bit**	**% A great deal**
Year 0 (*n*=41)	0	0	27	56	17
Year 1 (*n*=24)	0	0	36	50	13
Year 2 (*n*=20)	0	0	20	75	5
Year 3 (*n*=8)	0	0	13	63	25

Initial characterizations of Carl's and Kari's lessons. We used specific statements from Carl's and Kari's beliefs and video clip interviews to determine which teaching component each prioritized (Table 8). Carl and Kari both attended to student engagement, student efficacy, and concept development, however, they placed different degrees of emphases on these three components. Carl focused much of his planning, assessment, and teacher talk to support student engagement and building student efficacy. Conversely, while Kari also considered her students' engagement and efficacy, her classroom discourse revealed her stronger emphasis on concept development.

Table 8. *Initial characterization of Carl's and Kari's teaching goals*

	Teacher	**Theme**	**Example**
Student Engagement	Carl	PK	Learning theory: Uses discrepant events to generate interest. (CTBI_engage)
	Carl	Student readiness	Students need support applying math skills, such as graphing, in order to maintain engagement with science content (CVCI_EOS#1).
	Carl	SMK	Comfort with content allows Carl to focus on planning for and attending to classroom management. (CFUI_EXP/SMK#1)
	Kari	PK	"There's so much abstract thought…I try to get more tangible [examples] they can actually grasp" (KTBI_EOS#2)
	Kari	Class routines	"I want discussion, but I want it related. It's figuring out how loose you can let the reigns out before you have to bring it back in" (KTBI_OCT#3)

	Kari	PK	"The hardest part is if you have kids who aren't interested [in the topic], who just won't do it. It varies from class to class." (KTBI_VoTL#3)
Student Efficacy	Carl	Class routines	Developing classroom routines such as note taking, using notes, paying attention, working in groups (CTBI_EXP_SSE#1).
	Carl	Student resistance	Some students hesitate to engage in class content: "If I could convince him…that academics is something he wants to focus on, he'd do just fine" (CVCI_resist)
	Carl	Student readiness	"I have to spend a lot of time training that unspoken, implicit academic expectation" (CFUI_SR#1).
	Kari	PK	Balances difficulty level so students are appropriately challenged without "being overly frustrated, because then they'll just shut down" (KTBI_VoTL#3)
	Kari	Student support	"I think they need affirmations, because in lab they're always afraid they're going to screw something up." (KVCI_EOS#4)
	Kari	Student support	Reports her students are more engaged and confident when she is "next to them. They don't need me to explain it, but they want me there just in case." (KVCI_EOS#3)
Concept Development	Carl	Role as MKO	Models desired dispositions, establishes self "not in a position of absolute authority" but as a guide: "I know where we're going, come this way, you're taking a wrong turn" (CFUI_MKO).
	Carl	Assessment	Uses prepared rubric to make learning goals clear to students, uses "level 4 questions" to promote deeper group discussion
	Carl	Assessment	Three stage quiz cycle allows students to critique their own work for content accuracy and personal growth.
	Kari	Time for topic	"I gauge if we need an extra day to work problems, or explain reactions. We'll spend the extra time if we need." (KTBI_OCT#1)
	Kari	Making connections	"I will always stop and ask the 'why.' I think tying it all together and making them think about it is what my goal is." (KVCI_IG#1)
	Kari	Personal connection	Related a story of a current student who "had a light bulb moment. And he was, just, so loud!" to a similar experience as a GA: "I felt validated. It was good." (KTBI_EOS#5)

Kari's use of discourse to support concept development: "Getting to the why"

Kari teaches introductory chemistry, Advanced Placement chemistry, organic chemistry, and forensics at St. Sebastian High School. St. Sebastian is a private Catholic, all-boys boarding school, located on the rural edge of the state's largest urban center. St. Sebastian does not report student demographics to the State Education Agency, however, according to the school's promotional materials, St. Sebastian serves 225 students. About two-thirds of St. Sebastian's students live on campus; approximately 80% are Caucasian, and about 15% are visiting foreign students. Approximately 20% of St. Sebastian's students receive tuition support in the form of financial aid, work-study scholarships, or grants. During the academic year of the study, St. Sebastian's tuition, including room and board, was $17,500. St. Sebastian is a college-preparatory school and nearly all (99%) St. Sebastian students enter college or university after graduation, and the school boasts an average 28.0 ACT score for its recent graduates. Kari is one of three science teachers at St. Sebastian, and one of her St. Sebastian science colleagues is a fellow MAT graduate.

Kari's classes met for forty minutes each day. In many of our discussions, Kari mentioned the difficulty of conducting a lab exercise and discussing the results with such short periods. To address this challenge Kari reported that she tried to schedule pre-lab briefings at the end of one class period, have her students conduct the lab exercise the next, and then follow up with discussion of results or concepts during the following day's class.

When I (Aaron) observed one of Kari's post-lab discussion days, Kari told me that while she reviewed a homework assignment she noticed her students were confusing the predictive rules for single-replacement and double-replacement reactions. Kari addressed the misapplication during the post-lab discussion as documented in Table 9.

Table 9. *Excerpts of Kari's classroom dialogue with teacher comments (Concept Development)*

		Classroom dialogue	Kari's comments while watching video
[1]	**Kari:**	So sodium is higher on the activity series than magnesium. Just looking at the activity series, what does that mean?	"I try to give a leading question that isn't giving away the answer."
[2]	**Student 4:**	They can't switch…	
[3]	**Kari:**	What can't switch?	
[4]	**Student 4:**	Magnesium and sodium.	
[5]	**Kari:**	Depends, on the situation. I'm looking at sodium here, magnesium here. What can you tell me about these two metals, in relation to one another, Student 5?	Kari reported her students were confused about single and double replacement on homework assignments.
[6]	**Student 5:**	Sodium can replace magnesium, but magnesium can't replace sodium.	
[7]	**Kari:**	Why?	
[8]	**Student 5:**	Because sodium is higher on the reactivity…on the list.	
[9]	**Kari:**	Which means what?	"Let's use everything to understand the 'why.' So sodium is more reactive. Why is sodium more reactive?"
[10]	**Student 5:**	It has higher reactivity.	
[10]	**Kari:**	OK. Which means what?	
[10]	**Kari:**	You guys are so close here. Student 6. [Calls on Student 6 to answer.]	
[10]	**Student 6:**	Sodium doesn't, sodium wants to move around more.	

[2]	**Kari:**	Wow, are we going all the way back to ionization energy?	I want to be sure they don't just memorize…that's not understanding chemistry, that's [using] a graph or a table.
[3]	**Students:**	Oh, I remember that…	

In support of her students' efficacy, Kari reported she perceived a need for "affirmations, because they're afraid they're going to screw something up." During my visits, I observed Kari's students were cooperative and generally engaged in the class, and nearly every student participated in the class activities, including completing most of each homework assignment. Kari identified her main instructional goal: "getting to the why," or helping her students understand and describe the theoretical or microscopic explanation of each event. Kari told me her students could readily describe an observed phenomenon, but struggled to explain the phenomenon, if they attempted an explanation at all.

Kari recognized that her sophomores often had difficulty visualizing abstract concepts, and her approach was to ask "questions to try to get them there, instead of just expecting them to figure it out on their own" (KVCI_IG_redoxlab).

> I want to be sure they don't just memorize 'this has to be higher than this in order for it to replace,' because that's not understanding chemistry. That's understanding how to use a graph or use a table. And in the grand scheme of things, you could make it through that way, but [since] we're already talking about [related chemical concepts], let's use everything to understand the 'why.' So sodium is more reactive. Why is sodium more reactive?

Carl: Using classroom discourse to "show them they can do it."

Carl teaches physics and physical science at Honeydew Magnet High School. Honeydew is one of seven high schools in Urban Public School District (UPSD), and, according to the State Educational Agency's report, has a high poverty (81%) and student mobility rate (33%). About 18% of Honeydew's students are English Learners, and about 10% have recently arrived as refugees. Honeydew is a majority minority school (74% minority), and Carl, like two-thirds of his colleagues at Honeydew and 90% of UPSD teachers, is White. We arranged to observe one of Carl's freshman physical science classes. Like Kari's Introductory Chemistry class at St. Sebastian, this is the first physical science class for students at Honeydew.

Planning with evaluation in mind. Carl bases his lesson and unit planning on his assessment plan, using main ideas and a performance rubric to guide the content, the time he allocates for each concept, and the class activities he plans:

> "To make the quizzes, structured the way they are, I have to go through and make a rubric. I have to decide which [concepts are fundamental concepts], which ones are the everyday application questions, and which ones are the hard ones. So I've got a rubric with all [these questions in mind]. So when [I'm teaching class] those are the questions I'm asking. It's not so much teaching to the test, which I try to avoid, but it's making sure what I'm teaching is what I'm going to assess them on."

Carl reported that during the previous school year, he constructed his quizzes according to the district standards, which required each test or quiz to have depth of knowledge (DOK) Level 2, Level 3, and Level 4 questions. Carl explained his understanding of the level of questioning. According to Carl, Level 2 questions assess basic knowledge, such as definitions and terms that were "simple, straight forward, such as 'speed equals distance divided by time." Level 3 questions are application questions "and every student is supposed to hit that level." Level 4 questions are "above and beyond," meaning that students are required to extend what was taught in class and apply it to real world situations. Every test was supposed to have all three levels of questions, and teachers were to evaluate the test results based upon the highest level question the student answered correctly. Additionally, in terms of formative assessment Carl uses the quiz results to help students recognize the results of their effort in class during the assessment period, as a way to help: (a) students reflect on their learning, (b) build student self-efficacy, and (c) promote self-regulation.

Building student efficacy. Carl reported that in his experience, his students learn best from

> "hands-on activities, with the concrete thinking questions does a pretty good job. But the moment I start asking them an abstract question, like 'where did the bubbles come from?' and I get a lot of 'IDK' written on there—'I don't know.'" (TBI interview)

and "taking notes doesn't do them any good. Abstract discussions usually don't go over very well." Carl prefers to "get them moving and then ask them the hard questions. That's the most fun I have and I think it's when they learn best."

The "hard questions" are the application and analysis questions Carl included in the rubric:

> I will come over, and agitate them a little bit…use those Level 4 questions. [The group might be] making the graph of mass vs force, so I might ask them 'how would you measure this without using gravity?' Since they're ninth graders, they might look at me and say 'gravity's up and down' and I might say 'what can you do side to side to get to the difference between the two masses?' And I'll leave that group alone to stew, and I'll move on to another group who's struggling with how to make the graph or how to not pull each other's hair. (TBI Interview, CTBI-EXP-PK#1)

Carl told me his students often hesitate to begin work or to engage in on-topic discussions, and during the lessons I observed, he often moved from group to group to keep students engaged in the content or to encourage students to start the activity. During the lab portion of this lesson, Carl worked with student groups initially to set up the activity and model measurement, and later to help them figure out the graph. I showed Carl a video clip of his interactions with a female student and asked him to narrate his decision-making process; this is summarized in Table 10.

Table 10. *Video excerpts of classroom dialogue with teacher comments*

		Classroom dialogue	Carl's comments on video of lesson
[1]	**Carl:**	What I need you to do for me is to make a graph. Do you remember how to make a line graph? [Carl walks with the student to her lab station.] Do you have a pen? So you're going to… [Carl explains the graphing procedure]…then you	In this case, it was mostly making sure the math wasn't the challenge. Because she was getting stuck on how to graph, and if you don't have that, it's hard to see the relationship. That's what a

		draw a straight line. It's not hard! Alright, so this is my time, and you did time at zero, and at two, and four…	lot of this help was on, how do you set up a graph, how high do you need to go, what's the scale? It's one of those things a lot of our students struggle with.
[2]	**Student:**	[Nods, starts work.]	
[3]	**Carl:**	It's like counting quarters. At two seconds, it's at one hundred. Wow, that guy's fast. At four seconds, geeze, he's going to be way up here!	Even when she has the right answers, she's one that wants that "OK, you got that right" and off she goes. Even when she's doing it right, she can't see when she's doing it right. It's kind of hard to tell if [she needs] affirmation or really doesn't, [or] can't evaluate her work.

In a separate interview immediately after class, I asked Carl about this interchange. Carl told me the student had created an accurate graph in a non-standard style, and that he took time to see if the graph made sense.

> If they had accurately plotted the data, I didn't want to undermine the work they had done. It was a perfectly fine stacked bar graph, it just wasn't what I was looking for. I'm not going to say it was wrong, because it's not. We've got a different way. (CARL OBS CI_11112014_CI)

I showed Carl a video clip of the interaction, and I asked Carl how he knows if this student needs affirmation or if she needs help developing the concept:

> Usually if she asks me a more complex question, she just needs the affirmation. If she comes up with a detailed question, then [I say] "yeah, that's more or less what we're looking for." If she comes up and says "how do I graph this" then I know that she's just lost. (CVCI CLIP #1)

I wondered why Carl did not say "graphing is something science students struggle with" or something more general about the difficulties of teaching graphing skills to ninth-graders; when he spoke, Carl placed an equal emphasis on "graphing" and "our students." Carl explained his own experience as a science student and his more recent experience as a student teacher contrasted with his current experience at Honeydew. Carl told me he was a quick study in math and physics as a student, and the students at the affluent and suburban school where he completed his student teaching internship were more worried about the "details of graphing, not the actual process." Carl told me his current physical science students, like their peers at Honeydew, had difficulty with many math concepts, and graphing was one of the more difficult skills for them to master.

Supporting student engagement. Carl invested much of his instructional time in class encouraging students to begin work, supporting student work in progress, or managing student behavior. Promoting student engagement is an instructional goal of Carl's; he explicitly mentioned his efforts to keep students engaged in all three interviews. Carl and I had a conversation after I observed one of his lessons, and I wrote in my memo for the day that Carl was aware of a finding in Honeydew's accreditation report that Honeydew students were compliant, but not engaged. I observed many in-class interactions where Carl offered supports to promote his students' engagement.

According to Carl, his 4A Physical Science students are caught in a causality loop; many students lack the confidence to engage with concepts, and by not engaging, they miss opportunities to grow in their efficacy. Carl told me he observed a lack of productive engagement and the associated lost opportunities for conceptual development, which led him to readjust his teaching to focus on developing confidence, promoting efficacy, and explicitly showing his students the connections between their effort and their achievement. He restructured his assessment practices around two interconnected purposes, building his students' confidence to attempt an unfamiliar task, and convincing his students that failure is an opportunity to grow and learn. In support of the growth mindset, Carl placed two small posters above the classroom door; one reads "Fail, Fail again, Fail better," the other, a Richard Feynman quote: "We are trying to prove ourselves wrong as quickly as possible, because only in that way can we find progress."

During a lesson on Newton's Third Law, Carl distributed spring scales to pairs of students, and directed them to pull against each other, and to adjust so one spring measured a different force than the other (Table 11).

Table 11. Excerpts of classroom dialogue with teacher comments

		Classroom dialogue	Carl's comments on video clip
[1]	**Carl:**	OK, who's going to pull 5? OK, you're going to pull 5 and you're going to pull 15.	
[2]	**Student 1 and 2:**	[*Pairs of students pull on spring scales.*]	
[3]	**Carl:**	[*To Student 1*] Look at yours. Get it down to five.	
[4]	**Student 1:**	[*Adjusts spring scale*]	
[5]	**Carl:**	[*To student 2*] OK, look at yours. Get it up to fifteen. When you get it figured out let me know.	
[6]	**Student 2:**	[*Starts to pull on spring scale*]	
[7]	**Carl:**	[*After several minutes, asks for attention of the class*] How many people could do [that]? Raise your hand if you could pull the same? [*Waits for response; there are no raised hands.]*	"I need to get better at [identifying] the response I'm looking for…I was getting a lot of visual nods or [hand] gestures meaning 'yes, no, eeesh.'"

[8]	**Carl:**	How many of you could pull entirely different numbers? You guys think it's impossible? It is. I like the words Kyle put to it earlier 'I can't because one controls the other.' Did you guys notice that? The one pulling harder set the other one? And the person pulling a little bit? They tried to lighten up and what happened?	"A lot of times I stick Kyle next to people who are fairly social but do need the help." Carl explains he uses peer-to-peer talk and considers student ability level, social tendencies, and individual student interactions when he assigns seats.
[9]	**Student 3:**	It was the same.	
[10]	**Carl:**	Yeah, it was always the same. Doesn't matter who's pulling. Do you remember who's bigger?	
[10]	**Multiple students:**	No…who's stronger…	
[10]	**Carl:**	It didn't matter, for anything. It doesn't matter up or down, left or right. The only way to do it is if you find a way to cheat the scales.	"I want them to [ask] 'what's going on' rather than [saying] 'that's what he said'…so we can say 'you saw this, does it match with what you think should happen?'"

Carl's classroom interactions with students were often focused on encouraging them to start a problem or investigation, and helping them persist in the procedure or in finding a solution. Carl told me he did not fault students for not trying if they truly believed they would not be successful.

> If they really don't think they have the capability to do it, and they don't try...it's not license for me to give up on them, but I can't really get angry with them for not trying. If they really don't think they can do it, then they're not even going to give it a shot.

Carl told me his guiding question is "*how do I show them they can do it?*" He acknowledged he would also think "*if you would only try*," but told me he would not allow himself to direct the thought toward his students. He explained "unless you...start showing them other ways where they are making progress," asking his students to begin a novel or unfamiliar task would be "like asking a fish to fly."

Discussion

The results of this study serve as a product starting point to think about the link between the new science curriculum and discourse practices in the classroom. As shown in the previous section, only about a third of the observed lessons involved students-led interactions while two-thirds of the lessons were delivered through the traditional IRE model. However, there is no evidence to conclude that one is less effective than the other. The emphasis on discourse does not require "a substitution of nontraditional for traditional lessons" but "a repertoire of lesson structures and teaching styles, and the understanding of when one or another will be most appropriate for an increasingly complex set of educational objectives" (Cazden, 2001, p.56) on the teacher's part. Our class activities coding does indicate that factors such as the subject, the lesson topic, the students, the time of instruction (regular/block), or the school culture to name just a few, may combine to influence the teacher's choice of discourse practices. It is also worth pointing out that research on classroom discourse has shown that authentic questions do not necessarily lead to more classroom interaction or more effective teaching; "inauthentic" questions soliciting brief answers could also turn a monolog into a dialogue (Cazden, 2001. P.46). As teacher educators, it is critically important to help science teachers find a balance between asking good questions and boosting student-centered interaction to help students to become "fluent speaker of science" (Lemke, 1990).

Paper #4: Assessment Practices of New Science Teachers

Results

Focusing on the assessment factors scale using the five items on the EQUIP, we found that in this induction phase, more science lessons scored at "proficient" or "exemplary" levels of inquiry on the *conceptual development* and *assessment type* items. However, lessons that were scored as "developing" or "pre-inquiry" were more common on the *prior knowledge assessment*, *student reflection*, and *role of assessing* items (Table 12).

Table 12. *Effective Aspects of Assessment: Percentage of Observed Science Lessons at "Proficient" or "Exemplary" Levels of Inquiry* (n=319 lessons, not teachers)*

Assessment Factor	**Student Teaching % (*n*=71)**	**Year 1 %** (*n*=116)	**Year 2 %** (*n*=95)	**Year 3 %** (*n*=37)	**Mean %** (with student teaching)	**Induction (Years 1-3) Mean %**
A1: Prior Knowledge	1	3	9	3	4	5
A2: Conceptual Development	24	18	43	53	34	38
A3: Student Reflection	6	3	15	0	6	6
A4: Assessment Type	24	15	32	45	29	31
A5: Role of Assessing	3	6	13	16	9	12

* Note: *Proficient scored a "3" and "Exemplary" scored a "4" on the EQUIP instrument.*

Teachers used questioning as part of the *role of assessing*, especially as whole group discussions, as a common strategy to assess students' understanding. Most of them uses questions that require little explanation. Frequently, new teachers implement IRE patterns to assess students during instruction. For example, John, a first-year chemistry teacher worked with 11th grade students, reviewing how to name ionic compounds.

> *(JT/Nov 24th, 2014)*
> *7:44 The teacher gives some instructions about using their textbooks and the page where they were working last Friday. John reminds the students the rules of how to name ionic compounds.* ***He has some compounds in the presentation. The students are going to***

> ***write the formulas. John shows how to "criss-cross" the numbers of the charges to build ionic formulas****.*
>
> *7:52* ***John is in the front of the classroom, asking the name of the compounds in the exercise. The students are giving the answers****.*

John uses his students' answers to check if the students could apply the rules he had explained before to this practice problems. But, John does not ask for explanations or justification of the students' answers. There is no connection with students' background or a real context. This practice of asking students' questions to solve problems is common among these science teachers. Another example is Steven, a biology teacher. In his first year, he showed a video about the digestive system to his 7th grade students. He asked them to list the parts of this system and label a drawing after the video.

> *(SP/May 14th, 2014)*
> *0:00 The students are sitting at their desks. The teacher is at the front of the room. Students will do a brainstorm of the digestive system after the video****. The students are listing some parts of the digestive system. The teacher writes on the board the students' answers****. Now the teacher draws a sketch of the digestive system and questions students as to where each part goes. Students say the parts and where to put each of them (Steven labels the drawing).* ***The teacher does what the students say. The teacher asks the students to not write down what he is doing****.*
>
> *4:11. They are still ordering the parts in the sketch of the digestive system. The teacher writes on the board at the front, and the students help give the answers. The teacher explains about the esophagus****. The teacher explains that their digestive system is below their ribs. He asks them to touch their sternum and shows them where.*** *They touch their sternums.*

Because it is work in progress, Steven required the students to not write down what he was co-constructing on the board yet. He asked information about what the students can remember after the video and some other information that the students had before that lesson about the digestive system. After that, Steven directed and completed the parts of the digestive system. In these two examples, we observe how teachers assessed students' understanding through using questions and whole group discussion. John asked students to name compounds and Steven used brainstorming after the video to inform their teaching. In Steven's example, we can see some use of this information to adapt his explanation. Nevertheless, we see no explicit adaptations of his original plan, however this may be because we did not have access to it.

On Table 13 we summarize teachers' self-efficacy, specifically inquiring as to how much teachers believe that they can adjust lessons to proper level for individual students. In Year 1 and 2, 50% of teachers considered they have "some" skills and knowledge. By Year 3, the percentage of "some" increased to 62%, but this also mean that there were fewer teachers who believed that they could do more. For new teachers, adapting their curriculum and activities to different students' need is a challenge. Through assessment practices teachers are challenged to adapt their lessons based on the students' needs for learning in an inquiry lesson. Teacher educators and administrators could consider this information when they prepare and work with new teachers.

Table 13. *Teacher responses to Question* 17: How much can you do to adjust your lessons to the proper level for individual students?
(*n= teachers who completed survey at the end of teaching year*)

	Nothing (%)	**Very little (%)**	**Some (%)**	**Quite a bit (%)**	**A great deal (%)**
Year 1 (*n*=24)	0	0	50	46	4
Year 2 (*n*=20)	0	0	50	35	15
Year 3 (*n*=8)	0	0	62	25	13

We observed that on average 12% of lessons by the MAT teachers used explicit adaptations or lesson plan modifications after questioning or more formal types of assessments. This does not mean teachers did not modify their plans as part of this assessment for learning. But, we do not know how teachers makes these decisions without access to their original lesson plans. Due to the limits on the number of classroom observations we conducted annually, we may not have observed how teachers followed through using such information about learners.

We did observe some teachers using questions to assess understanding. For example, some were more concerned about analyzing students' answers than simply "right" or "wrong" answers. For example, Nick, a third-year 6th grade teacher at the beginning of class:

> *(NM/October 2nd, 2014)*
> *8:58 Nick is showing the directions written on the slide about the bell work. He explains they are going to work on their notebooks. They are going to write the definition of a biome.* ***"Write down what do you think. Don't worry if you are wrong or right"****. Then, Nick explains, they are going to write down the definition of the book, and compare both definitions. A student says it is hard to come up with a definition. Students are working. Nick gives the page where the book's definition is. He is lending books to students who do not have their books. The second question is what biome (the town where the school is located) is within. Nick waits for the students to finish.*

Nick's example describes a strategy to assess understanding. There are many others. Evaluating "right" and "wrong" answers is still a common practice in the science classroom for these teachers. Like in Steven's and John's examples, questioning was still about finding the answer to a particular problem. Teachers seem to be focused on content knowledge acquisition, and not that much on students' thinking processes, as it is reported in other studies (Coffey, Hammer, Levin, & Grant, 2011).

Through this exploratory study we observed that teachers also grew in their self-efficacy about implementing alternative teaching strategies in their classrooms. On their third year, 62.5% of teachers answered "quite a bit" versus the 33% of teachers in their first year (Table 14). The need to implement an alternative strategy could come from the assessment for learning. It seems that the ability to assess understanding and use alternative strategies to modify an original plan inside the science classroom is something that teachers improve as they gain experience.

Table 14. *Teacher responses to Question 23*: How well can you implement alternative strategies in your classroom?
(*n*= *teachers who completed survey at the end of teaching year*)

	Nothing (%)	**Very little (%)**	**Some (%)**	**Quite a bit (%)**	**A great deal (%)**
Year 1 (*n*=24)	4	0	58	33	4
Year 2 (*n*=20)	0	0	45	50	5
Year 3 (*n*=8)	0	0	37.5	62.5	0

The types of activities and their use of higher thinking skills (HOS) and critical thinking, teachers used grew in complexity over time. This is suggested by the increased percent in the *conceptual development* item score during teachers' induction phase. In 53% of Year 3 teachers' lessons, they used open-ended questions and data analysis for assessment for learning purposes. There was a constant increase from Year 1 to Year 3 teachers (18% to 53%).

We rarely observed formal and informal assessment practices based on argumentation (e.g., evidence to support claims) and the connection between different concepts. However, learning strategies using repetition and memorization were also not commonly present in the lessons we observed. This may indicate that some of the reform-based practices that were promoted in the teacher education program persisted in teachers' long-term behaviors. However, this will need to be investigated further. The growth in students' *conceptual development* was consistent with the *assessment type* used during these lessons. In 31% of lessons, teachers used authentic measures in formal and informal assessment activities. Growth was almost 15% every year from teachers' first to third year teaching (15% to 45%). In the lessons we observed, teachers still used factual and discrete knowledge to assess understanding, but they began to incorporate more authentic measures into their assessment practices. This is also directly related to reform-based practices and the development of students' critical thinking.

We observed that teachers assessed students' prior knowledge and modified their instruction about 5% of the time. It was common for teachers to assess students' understanding as a review of what they studied in previous lessons. "Bellwork" or "warm-up" questions at the beginning of the class about the previous session are a common practice. Nevertheless, our observations of prior knowledge assessment, based on students' background or experiences about particular topics were rare. Here is an example of a prior knowledge exploration from Charlotte, a physics teacher talking about gravity with her 11th grade students during her first year of teaching:

> (CR/Jan 27th, 2014)
> *12:33 Charlotte is still lecturing. She is asking if they have been in the ocean. Tomorrow they will talk about tides. She asks for them to take out a piece of paper* ***. In one or two sentences she asks the students to write what they know about tides. She tells them that she is not going to grade it. She explains to the students that she wants to know their previous knowledge. Charlotte also asks them to write down what they want to know about tides.*** *She tells the students to hand in their answers about tides as well as the lab.*

By using a KWL assessment strategy Charlotte learns more about her students' ideas and how she might need to adjust her lesson. Overall, for most of the teachers we observed assessing prior knowledge appears to be an on-going challenge during their induction phase. We expect to observe more of this practice in the future, but we wonder how much professional development teachers may need to change this particular practice.

Student reflection was the other least observed assessment aspect. In only 6% of lessons observed did new science teachers explicitly encourage students to reflect on their learning at an understanding level or concerning their higher level thinking skills (e.g., to evaluate, to design, to predict). Reflection is an essential element for learning. It is strongly related to teachers sharing learning responsibility and students' empowerment (Dimick, 2012). It is probable that *student* *reflection* represents a real challenge for teachers in the induction phase as they did not appear to use it much in the lessons we observed.

In summary, these MAT induction teachers were in the practice of assessing students' understanding, and showed a consistent increase in the use of authentic measures and learning activities using CAT to assess their students. Nevertheless, predominantly we see teachers using factual and discrete knowledge to assess. Student reflection for the purpose of revealing understanding and improved metacognition is almost absent from these science teachers' lessons. Assessment and use of prior knowledge, metacognition and the use of alternative strategies during instruction are challenges for new teachers.

Discussion: Assessment

Although assessment for learning is a good teaching practice to increase students' learning, motivation, and self-esteem (Bookhard, 2009; Black and Wiliam, 1998), the results of this study also show us that it is a difficult task, especially for new teachers (Bell & Cowie, 2001). Nevertheless, we could see growth and incremental change in some of the constructs of assessment for learning as measured by the EQUIP.

For example, the EQUIP scale shows an incremental change in the *conceptual development* of the assessment practices and activities teachers use as part of their instruction. Assessment for CAT requires a shift in the assessment conception. Assessment should be considered more than a testing system to provide grades to students, especially for struggling learners. Brown, Afflerbach, and Croninger (2014) suggest that the assessment for learning must be based on learning progressions, real contexts where students apply in real life their performances, and feedback based on clear rubrics with excepted performances. Teachers still require working on involving connections with real contexts and other concepts when they use

CAT activities. Also, the use of authentic measures and real-life problems can contribute to improving this conceptual development. Professional development on strategies like problem-based learning and learning progressions can be a way for these teachers to think about how to increase their assessment for learning in an inquiry-based science lesson.

It might be helpful for these new teachers to use more local context (e.g., school and the students' interests) in their science instruction. Assessment of students' prior knowledge and reflection were almost absent from the lessons we observed. Although science education research recognizes the critical role that prior knowledge and metacognition plays within learning, there is a known gap between theory and practice. For example, metacognition is not a regular practice in science classrooms (Ben-David & Orion, 2012). Similarly, we rarely found good examples of prior knowledge assessment. Teachers need be open to learning during all their career, but especially during those first years of teaching (Luft, 2011). With some targeted mentoring, professional development, and reflection, teachers can learn and practice ways to apply them more frequently these assessment practices in their classrooms.

One important element that needs to be considered is assessment policies and practices in each of teachers' schools; in other words, the school culture. Teachers will align their practices and curriculum to what they are asked to do. Schools and teachers invest time in what is going to be evaluated (Berliner, 2010). How much space and promotion of scientific inquiry practices do teachers have in their schools? In what extent in this era of accountability large-scale testing has an impact on these new science teachers' assessment practices? Anderson (2012) explained that "teachers and administrators repeatedly expressed the feeling that accountability-based reform disrupts research-based reform efforts in science. They asserted accountability limits time and effort spent on science, drives the remaining science instruction toward memorization of facts, and constraints student learning" (p.121). Educational policies very often influences science education by prioritizing “some conditions of practice over others and emphasizing particular aspects of what and how of science teaching and learning” (Fensham, 2009, p. 1077).

Assessment for learning is a desirable practice in the science classroom. Science teacher education programs should provide preservice teachers with the knowledge and tools to help them to use it in their future classrooms. Educational policies and the school context should also provide an adequate environment for in-service teachers to use assessment effectively. More studies should be conducted to analyze the impact of these policies in new teachers’ classrooms. We also recommend further analysis of lesson plans, interviews, and more classroom observations to understand how different groups (e.g., middle vs. high school, high SES vs. low SES schools, in-field vs. out-of-field teachers) teachers may modify instruction after assessment practices.

Paper #5: Curricular Choices of New Science Teachers

Results

Analysis of the 319 science lessons generated from four cohorts of student teachers and graduates from the teacher education program showed the same general pattern of improvement in all four curriculum factors constructs (Table 15). There was a greater similarity between student teaching and Year 1 teaching percentages of effective teaching than in Years 2 and 3, all four scored items were 25% or less among the 71 student teaching and 117 Year 1 lessons that were analyzed. While *content depth* and *integration of content and investigation* showed greater improvement (from 24% to 47-51% and 34-45% respectively) from both student teaching and Year 1 lessons to Year 2 (*n*=93) and Year 3 (*n*=38) lessons, *learner centrality* (from 11% to 17-24%) and greater opportunities for students to *organize and record information* (from 6% to 14-

21%) showed less movement over the induction period. In other words, it appears that induction-level teachers provided more depth of content and better integration among content and investigation as teachers became more experienced, but that more opportunities could be made available for students to be more centrally-located within activities and have greater executive control over the ways in which they manage scientific information.

Table 15. *Effective Aspects of Curriculum: Percentage of Observed Science Lessons at "Proficient" or "Exemplary" Levels of Inquiry* (n=319 lessons, not teachers)*

Curriculum Factor	**Student Teaching % (*n*=71)**	**Year 1 %** (*n*=116)	**Year 2 %** (*n*=95)	**Year 3 %** (*n*=37)	**Mean %** (with student teaching)	**Induction (Years 1-3) Mean %**
C1: Content Depth	24	25	51	47	37	41
C2: Learner Centrality	11	17	17	24	17	19
C3: Integration of Content & Investigation	24	26	34	45	32	35
C4: Organization & Recording Information	6	9	14	21	13	15

* Note*: Proficient scored a "3" and "Exemplary" scored a "4" on the EQUIP instrument.*

In examining our survey data on teachers' self-efficacy for items that related to curricular factors, we focused on one item in particular, *#8: How well can you establish routines to keep activities running smoothly?* In response to this item, 78% of Year 1, 100% of Year 2, and 75% of Year 3 teachers indicated that they thought they could do "quite a bit" or "a great deal" to establish productive routines to support activities. Thus, we might expect teachers to feel highly efficacious in implementing learning activities with their students. However, when we review the observation data we do not see a high degree of learner centrality or even many opportunities for students to organize or record data. It may be that teachers' self-efficacy in establishing routines is insufficient for implementing more inquiry-based curriculum. It would be more helpful if the survey question could help distinguish between lower and higher levels of inquiry-based instruction. This is something that we could explore further through interviews about curricular choices in order to offer additional insights to the connections between teaching self-efficacy and learner-centered curriculum.

Interpretation

Two of the four curriculum items on the EQUIP instrument showed less change across the induction period; this indicates that some aspects of curriculum factors were more dominated by teacher activity. The lesser degree of learner centrality in which students were allowed to design aspects of their investigations and be more active participants in their learning during a science lesson reflects a limit to the degree to which teachers have integrated activities that allow for more than predictable results (i.e., verification-level labs). This is despite the fact that the 5E inquiry-based model was required of all preservice teachers in their lesson and unit plan designs in the MAT program. As indicated in *How Students Learn* (NRC, 2005) students need to learn executive control and self-regulation of their learning activities. Teacher professional development activities could potentially help teachers think differently about how to frame curriculum in ways that provide more opportunities to support students' development in these areas. Certainly problem- and project-based learning could expand students' opportunities to learn in more scientifically authentic ways.

Limitations of Study

There are several limitations of the study. First, this investigation was exploratory as we followed the first few graduate cohorts from the same MAT program and we expanded our research methods as time went on and particular questions about teachers' practice occurred to us. Also because we only have one cohort's (*n*=8 teachers) lessons that represented Year 3 of the induction period the study is still as yet underpowered to be able to employ modern statistical model building such as a hierarchical linear model . Second, we have been collecting another set of observed and coded lessons over the 2015-2016 academic year (n=225+) as well as interviewing teachers for additional lessons that surround the observed lesson. Thus, we hope to be able to more accurately characterize teachers' enacted practices as well as capture their perspective on what parts of their lessons they think are effective and which ones they want to change. We anticipate that our findings may change when we add more science lessons, thus the new data may confirm or refute general trends that we have seen in this initial sample of lessons.

Conclusions

Few studies have followed science teachers from their teacher education program into the field with this many teachers. Most available research about teachers emerging practices is in the form of case studies (Crawford, 2014). This 3-year longitudinal study describes the emergent practices from a single teacher education program, but its findings are transferable to other similar MAT programs that recruit teacher candidates with bachelor's degrees in science. Science teacher educators and professional development providers may find our results useful in thinking about teacher preparation priorities and induction phase teacher professional development needs. These findings can provide insights into issues new teachers face during their induction phase and the type of support they need to expand their teaching repertoire. We describe the direction in which this exploratory work is headed and our next steps to build more robust recommendations.

Future Work

The data from this study will be used along with the current year's (2015-2016) data set to build a hierarchical linear model of teacher change and use of inquiry-based instruction. We are adding another survey about teachers' beliefs about reform-based science teaching that will be accompanied by interviews. We are also tracking specific instructional strategies across time to determine which ones teachers use most and least frequently. To date we have not disaggregated the science lesson data by teachers' in-field content status, school level (i.e., middle and high school), or by socioeconomic status, all of which may influence the degree to which inquiry-based instruction may be used by beginning science teachers. Our new NSF Noyce grant (Track I, Phase II) is also currently funding a comparison study of teachers who have graduated from our undergraduate program. A few of the differences between these groups include that the undergraduate teachers complete their teacher licensure with less than an undergraduate degree in science and no teacher action research capstone project. We will be investigating if, by comparison, the MAT program accelerates new science teachers' growth or if there is more variance within groups then there is between them. This comparison will allow us to provide more specific recommendations to other teacher education programs and improve our science teacher preparation efforts.

To further support our findings, we are analyzing how these new science teachers' enacted curriculum and their self-efficacy influence specific curriculum and instruction practices in science classrooms. For instance, assuming that higher levels of inquiry would involve students more often in investigations and various activities requiring active engagement, we will examine the frequency of use of specific activities by teachers. To do this, we have already identified 45 classroom practices including opening engagement activities and prior knowledge assessment, different types of lab activities (e.g., verification, guided, and open inquiry), video, teacher-led (e.g., lecture and class demonstration), student-led (e.g., using technology, collecting data, and small group discussion), and classroom organization. Using our list of classroom activities, we have completed coding our field notes on these same 319 science lessons that we have coded using the EQUIP instrument and written about in this proposal. We will triangulate prevalent classroom activities across groups from preservice to Year 3 teachers as we continue to gather longitudinal data about the graduates of our MAT program.

Acknowledgements

This work was funded by a National Science Foundation grant, NSF # 1540797.

References

Anderson, K. J. B. (2012). Science education and test-based accountability: Reviewing their relationship and exploring implications for future policy. *Science Education*, *96*(1), 104–129.

Anderson, K. T., Zuiker, S. J., Taasoobshirazi, G., & Hickey, D. T. (2007). Classroom Discourse as a Tool to Enhance Formative Assessment and Practice in Science. *International Journal of Science Education*, *29*(14), 1721–1744.

Banilower, E.R., Smith, P.S., Weiss, I.R., Malzahn, K.A., Campbell, K.M. and Weis, A.M. (2013). *Report of the 2012 National Survey of Science and Mathematics Education.* Chapel Hill, NC: Horizon Research.

Banilower, E.R., Trygstad, P.J., and Smith, P.S. (2015). The First Five Years: What the*2012 National Survey of Science and Mathematics Education* Reveals about Novice Science Teachers and their Teaching. In J.A.Luft and S.L. Dubois (Eds.), Newly Hired Teachers of Science, 3-29.

Basu, S.J., & Calabrese Barton, A. (2007). Developing a sustained interest in science among urban minority youth. *Journal of research in science teaching*. 44 (3), 466-489.

Bell, B. (2007). Classroom assessment of science learning. In S.K. Abell and N. Lederman (Eds.), *Handbook of Research on Science Education* (pp. 965-1006). Mahwah, NJ: Lawrence Erlbaum.

Bell, B., & Cowie, B. (2001). *Formative Assessment and Science Education.* The Netherlands: Kluwer Academic Publishers.

Ben-David, A., & Orion, N. (2012). Teachers' Voices on Integrating Metacognition into Science Education. *International Journal of Science Education*, 35(18), 1–33.

Berliner, D.C. (2010). The incompatibility of high-stakes testing and the development of skills for the twenty-first century. In R. Marzano (Ed.), *On Excellence in Teaching* (pp. 112-143). Bloomington, IN: Solution Tree Press.

Bianchini, J. (2012). Teaching While Still Learning to Teach: Beginning Science Teachers' Views, Experiences, and Classroom Practices. In B.J. Fraser, K. Tobin, & C.J. McRobbie (Eds.). *Second International Handbook of Science Education* (pp. 389-399). The Netherlands: Springer.

Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *The Phi Delta Kappan*, 80 (2), 139-144, 146-148.

Blanchard, M. R., Southerland, S. a., Osborne, J. W., Sampson, V. D., Annetta, L. a., & Granger, E. M. (2010). Is inquiry possible in light of accountability?: A quantitative comparison of the relative effectiveness of guided inquiry and verification laboratory instruction. *Science Education*, 94(4), 577–616.

Brookhart, S. M. (2009). Assessment and Examinations. In *International Handbook of Research on Teachers and Teaching* (pp. 723-738). Springer US.

Brown, N. J. S., Afflerbach, P. P., & Croninger, R. G. (2014). Assessment of Critical-Analytic Thinking. *Educational Psychology Review*, 26(4), 543–560.

Brown, P., Friedrichsen, P., & Abell, S. (2012). The Development of Prospective Secondary Biology Teachers PCK. *Journal of Science Teacher Education*, 24(1), 133–155.

Bryan, L. (2012). Research on Science Teacher Beliefs. In B.J. Fraser, K. Tobin, & C.J. McRobbie (Eds.), *Second International Handbook of Science Education* (pp. 477-495). The Netherlands: Springer.

Cazden (2001). *Classroom discourse: The language of teaching and learning.* Portsmouth: Heinemann.

Coffey, J. E., Hammer, D., Levin, D. M., & Grant, T. (2011). The missing disciplinary substance of formative assessment. J*ournal of Research in Science Teaching*, 48(10), 1109–1136.

Crawford, B. (2007). Learning to teach science as inquiry in the rough and tumble of practice. *Journal of Research in Science Teaching, 44*(4), 613-642.

Darling-Hammond, L., & Bransford, J. (Eds.). (2007). *Preparing Teachers for a Changing World: What Teachers Should Learn and Be Able to Do*. San Francisco, CA: Jossey-Bass, A Wiley Imprint.

Davis, E. A., Petish, D., & Smithey, J. (2006). Challenges new science teachers face. *Review of Educational Research, 76*(4), 607-651.

Deboer, G. (2014). The History of Science Curriculum Reform in the United States. In S. K. Abell & N. G. Norman (Eds.), *Handbook of Research on Science Education Vol. II* (559-578). New York, NY: Routledge.

Dewey, J. (1910). *How we think*. United States of America: D.C. Heath & Co., Publishers.

Dimick, A. S. (2012) Student empowerment in an environmental science classroom: toward a framework for social justice science education. *Science education*. 96(6), 990-1012.

Duit, R., & Treagust, D. F. (2003). Conceptual change: A powerful framework for improving science teaching and learning. *International Journal of Science Education*, 25(6), 671–688.

Edwards, G., & Thomas, G. (2010). Can reflective practice be taught? *Educational Studies*, 36(4), 403–414.

Evans, B. (2011). Content knowledge, attitudes, and self-efficacy in the mathematics New York City Teaching Fellows (NYCTF) Program. *School Science and Mathematics, 111*(5), 225-235.

Fensham, P. J. (2009). The link between policy and practice in science education: The role of research. *Science Education*, *93*(6), 1076–1095.

Furtak, E. M., Morrison, D., & Kroog, H. (2014). Investigating the Link between Learning Progressions and Classroom Assessment. *Science Education*, 98(4), 640–673.

Gee (2014). *An introduction to discourse analysis: Theory and method.* New York: Routledge.

Guskey, T. (1988). Teacher efficacy, self-concept, and attitudes toward the implementation of instructional innovation. *Teaching and Teacher Education, 4*(1), 63-69.

Haag, S., & Megowan, C. (2015). Next Generation Science Standards: A national mixed-methods study on teacher readiness. *School Science and Mathematics*, 115(8), 416-426.

Haug, B. S., & Ødegaard, M. (2015). Formative Assessment and Teachers' Sensitivity to Student Responses**.** *International Journal of Science Education*, *37*(4), 629–654.

Kelly, G. J. (2008). *Learning science: Discursive practices.* In M. Martin-Jones, A. M. de Mejia and N. H. Hornberger (eds), Encyclopedia of Language and Education, 2nd Edition, Volume 3: Discourse and Education, 329–340.

Kim, B. (2001). Social constructivism. In M. Orey (Ed.), *Emerging perspectives on learning, teaching, and technology*. Available Website: http://www.coe.uga.edu/epltt/SocialConstructivism.htm

Kind, P. M. (2013). Conceptualizing the Science Curriculum: 40 Years of Developing Assessment Frameworks in Three Large-Scale Assessments. *Science Education*, *97*(5), 671-694.

Lee, O. (1999). Science knowledge, world views, and information sources in social and cultural contexts: Making sense after a natural disaster. *American Educational Research Journal*, *36*(2), 187-219.

Lee, O. (2013). Oral Discourse in Teaching and Learning Science in Relation to the Next Generation Science Standards. Paper prepared for the National Research Council Conference on "Literacy for Science in the CCSS and NGSS". December 9 and 10, 2013 Washington, DC.

Lemke, J. L. (1990). *Talking science: language, learning and values*. Norwoord, NJ: Ablex.

Lewis, E.B., Musson, A., & Lu, J. (2014, March). *Educating highly-qualified science teachers: Challenges and perspectives*. Paper presented at the 2014 Annual International NARST Conference, Pittsburgh, Pennsylvania.

Lewis, E.B., Rivero, A., Musson, A., Lu, J., and Lucas, L. (2015, October). *Enacted teaching practices of induction-level science teachers*. Poster session presented at the 2015 Regional Meeting of the NSF Robert Noyce Program Conference, Omaha, Nebraska.

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).

Marshall, B., & Drummond, J. M. (2006). How teachers engage with Assessment for Learning: Lessons from the classroom. *Research Papers in Education*, *21*(2), 133–149.

Marshall, J. C., Horton, B., & Smart, J. (2009). 4E × 2 instructional model: Uniting three learning constructs to improve praxis in science and mathematics classrooms. *Journal of Science Teacher Education,* 20(6), 501–516.

Marshall, J.C., Horton, B., Smart, J, & Llewellyn, D. (2008). *EQUIP: Electronic Quality of Inquiry Protocol: Reviewed from Clemson University's Inquiry in Motion Institute*, www.clemson.edu/iim.

Marzano, R.J. (2010). *Formative assessment and standards-based grading*. United States of America: Marzano Research Laboratory.

McGinnis, J.R., Parker, P. and Graeber, A.O. (2004). A cultural perspective of the induction of five reform-minded beginning mathematics and science teachers. *Journal of Research in Science Teaching*, *41*(7), 720-747.

Mortimer, E., & Scott, P. (2003). *Meaning Making in Secondary Science Class- rooms*. Open University Press, Maidenhead (UK).

National Research Council (1996). *National science education standards: Observe, interact, change, learn*. Washington, D.C.: National Academy Press.

National Research Council (2005). *How students learn: History, mathematics and science in the classroom,* A Targeted Report for Teachers, M Donovan and J. Branford (Eds.). Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academy Press.

National Research Council (2010). *Preparing teachers: Building evidence for sound policy*. Washington, D.C.: National Academies Press.

NGSS Lead States (2013). Next Generation Science Standards: For States, By States. Achieve, Inc. on behalf of the twenty-six states and partners that collaborated on the NGSS.

Nilsson, P. (2013). What do we know and where do we go? Formative assessment in developing student teachers' professional learning of teaching science. *Teachers and Teaching*, *19*(2), 188–201.

Osborne, J. (2007). Science education for the twenty first century. *Eurasia Journal of Mathematics, Science and Technology Education*, 3(3), 173–184.

Osborne, J. (2010). Arguing to learn in science: the role of collaborative, critical discourse. *Science (New York, N.Y.)*, *328*(5977), 463–466.

Osborne, J. (2014) Scientific practices and inquiry in the science classroom. In N.G. Lederman & S.K. Abell (Eds.). *Handbook of research on science education* (pp. 579-620). New York: Routledge Taylor & Francis Group.

Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research, 62*(3), 307-332.

Pellegrino, J. W. (2012). Assessment of science learning: Living in interesting times. *Journal of Research in Science Teaching*, *49*(6), 831–841. http://doi.org/10.1002/tea.21032

Pintrich, P. R., Marx, R. W., & Boyle, R. a. (1993). Beyond Cold Conceptual Change: The Role of Motivational Beliefs and Classroom Contextual Factors in the Process of Conceptual Change. *Review of Educational Research*, 63(2), 167–199.

Roth, W. M. (1996). Teacher questioning in an open-inquiry learning environment: Interactions of context, content, and student responses. *Journal of Research in Science Teaching*, *33*(7), 709-736.

Scott, P., Asoko, H., Leach, J. (2007). Student conceptions and conceptual learning in science. In Abell, S. & Lederman N. (Eds.). *Handbook of Research on Science Education.* (pp. 31-56). New York: Routledge.

Siegel, M. a., & Wissehr, C. (2011). Preparing for the plunge: Preservice teachers' assessment literacy. *Journal of Science Teacher Education*, 22(4), 371–391.

Songer, N. B., & Ruiz-Primo, M. A. (2012). Assessment and science education: Our essential new priority? *Journal of Research in Science Teaching*, *49*(6), 683–690. http://doi.org/10.1002/tea.21033

Supovitz, J. A., Mayer, D. P., & Kahle, J. B. (2000). Promoting inquiry-based instructional practice: The longitudinal impact of professional development in the context of systemic reform. *Educational Policy*, 14, 331-356.

Sutton, C. (1996). Beliefs about science and beliefs about language. *International Journal of Science Education*, *18*(1), 1-18.

Treagust, D. (2010). General instructional methods and strategies. In S. K. Abell & N. G. Norman (Eds.), *Handbook of Research on Science Education* (373-391). New York, NY: Routledge.

Treagust, D. & Tsui, C. (2014). General instructional methods and strategies. In S. K. Abell & N. G. Norman (Eds.), *Handbook of Research on Science Education Vol. II* (303-320). New York, NY: Routledge.

Tschannen-Moran, M. & Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education, 17*, 783-805.

Wiggins, G. (1998). *Educative assessment, designing assessments to inform and improve student performance*. New Jersey: Jossey Bass.

Wiliam, D. (2007). Keeping learning on track: Classroom assessment and the regulation of learning. In F.K. Lester (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 1053-1098). Charlotte, NC: Information Age Publishing.

Wilson, M.R., & Bertenthal, M.W. (Eds.) Committee on Test Design for K-12 Science Achievement, (2005). *Systems for State Science Assessment.* Washington, D.C,: National Academies Press.

Woolfolk, A. & Margetts, K. (2007). *Educational psychology*. Frenchs Forest, N.S.W.: Pearson.

Zeidler, D.L. (2014). Socioscientific Issues and a Curriculum Emphasis. In S. K. Abell & N. G. Norman (Eds.), *Handbook of Research on Science Education Vol. II* (697-726). New York, NY: Routledge.

Zohar, A. & Barzilai, S. (2013). A review of research on metacognition in science education: current and future directions. *Studies in Science Education*, 49(2), 121–169.