TEACHER ENGAGEMENT PRACTICES IN INCLUSIVE MATHEMATICS CLASSROOMS: AN INDIRECT REPLICATION STUDY OF THE EFFECTS OF DIRECTED CONSULTATION

A Dissertation in
Special Education

by
Karen Rizzo

© 2018 Karen Rizzo

Submitted in Partial Fulfilment of the Requirements for the Degree of

Doctor of Philosophy

August 2018
The dissertation of Karen Rizzo was reviewed and approved* by the following:

David Lee  
Professor of Special Education  
Head of the Department of Educational Psychology, Counseling, and Special Education  
Dissertation Advisor  
Chair of Committee

Paul J. Riccomini  
Associate Professor of Special Education

Mary Catherine Scheeler  
Associate Professor of Special Education  
Professor-In-Charge

Lisa Kopp  
Associate Professor of Human Development

*Signatures are on file in the Graduate School
ABSTRACT

This study examines the effects of an approach to professional development known as the Supporting Early Adolescent Learning and Social Success (SEALS) framework. The intervention was designed to support teachers in improving the academic, social, and behavioral adjustment of middle school students in diverse classrooms. Online core and supplemental modules were completed by each teacher as well as directed consultation meetings which occurred, on average, once per week throughout the school year. As part of a larger data set, this study specifically examines the data from eighth grade mathematics teachers with inclusive classrooms. Teachers’ engagement practices for beginning and during-instruction were examined for changes following implementation of the professional development. In general, practices that include a match between instruction and high levels of on-task time lead to an increase in academic engagement of students (Gersten et al., 2009; Pashler et al., 2007). Using a single-case, multiple-baseline design, data were collected on the engagement practices of six teachers. Visual analysis of the data show mixed results with complementary small effects found in Tau-U calculations. Overall, results indicate SEALS made a small but targeted impact on individual teachers’ engagement practices in inclusive, eighth grade mathematics classrooms.

Keywords: Supporting Early Adolescent Learning and Social Success (SEALS), engagement, directed consultation (DC), professional development
TABLE OF CONTENTS

LIST OF TABLES .......................................................................................................................... viii
LIST OF FIGURES ......................................................................................................................... ix
ACKNOWLEDGEMENTS ............................................................................................................... x
INTRODUCTION ............................................................................................................................. 1
  Common Core Math Standards (CCMS) and Professional Development ......................... 2
  Engaging Middle School Math Teachers in Professional Development ......................... 3
  Focus and Delivery of Professional Development ................................................................. 4
Supporting Early Adolescent Learning and Social Success (SEALS) ................................... 6
  What is SEALS ............................................................................................................................ 6
  SEALS and Multi-Tiered Systems of Support for Students ..................................................... 7
  Components of SEALS ............................................................................................................. 8
  Previous SEALS Research ......................................................................................................... 9
Summary and Purpose of the Current Study ............................................................................. 11
Research Questions ..................................................................................................................... 12
METHOD ......................................................................................................................................... 12
  Participants and Setting ............................................................................................................. 12
  Procedures .................................................................................................................................. 13
    Pre-baseline ............................................................................................................................. 13
    Baseline .................................................................................................................................. 13
    Intervention .............................................................................................................................. 13
Dependent Variables and Data Collection ............................................................................. 13
  Design ......................................................................................................................................... 17
  Data Analysis ............................................................................................................................. 17
  Inter-observer Agreement ........................................................................................................ 18
  Intervention Fidelity .................................................................................................................. 18
RESULTS ......................................................................................................................................... 19
  Total Score ................................................................................................................................. 19
  Beginning-Instruction Activities ............................................................................................... 20
    Alignment of Instruction ......................................................................................................... 20
    Statement of Clear Behavioral Expectations ......................................................................... 21
    Materials Ready at the Start .................................................................................................... 22
    Opening/Readiness Activity ................................................................................................. 23
  During-Instruction Activities ..................................................................................................... 23
    Restatement of Behavioral Expectations ............................................................................... 23
    Maintenance of Time-On-Task .............................................................................................. 24
Frequent Review/Summarizing .................................................. 25
DISCUSSION.................................................................................. 26
  Intervention Effects ................................................................. 27
  Total Score ............................................................................. 27
  Beginning-Instruction ............................................................ 28
  During-Instruction ................................................................ 29
Previous Research on SEALS and Relationship with Present Work........... 29
Reform-based Professional Development ............................................. 30
Limitations and Future Research .................................................... 32
Implications for Practice ............................................................. 34
Conclusion .................................................................................. 34
REFERENCES .............................................................................. 36
TABLES AND FIGURES ................................................................ 43
  Table 1 .................................................................................. 43
  Table 2 .................................................................................. 44
  Table 3 .................................................................................. 46
  Table 4 .................................................................................. 47
  Table 5 .................................................................................. 48
  Figure 1 ................................................................................ 49
  Figure 2 ................................................................................ 50
  Figure 3 ................................................................................ 51
  Figure 4 ................................................................................ 52
  Figure 5 ................................................................................ 53
  Figure 6 ................................................................................ 54
  Figure 7 ................................................................................ 55
  Figure 8 ................................................................................ 56
APPENDIX A: Literature Review: Professional Development for Increasing Academic Engagement in Math and Science
  Title Page................................................................................ 57
  Abstract .................................................................................. 58
  Introduction ............................................................................. 59
    The Need for Professional Development in STEM ................. 60
    Effective Professional Development in STEM ...................... 61
    What STEM PD Should Include ........................................... 61
      Content focus .................................................................... 61
      Active learning .................................................................. 62
      Coherence ......................................................................... 62
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Limitations</td>
<td>85</td>
</tr>
<tr>
<td>Implications for Practice and Research</td>
<td>85</td>
</tr>
<tr>
<td>Conclusions</td>
<td>86</td>
</tr>
<tr>
<td>References</td>
<td>88</td>
</tr>
<tr>
<td>Tables</td>
<td>96</td>
</tr>
<tr>
<td>Table 1</td>
<td>96</td>
</tr>
<tr>
<td>Table 2</td>
<td>100</td>
</tr>
<tr>
<td>Table 3</td>
<td>102</td>
</tr>
<tr>
<td>Table 4</td>
<td>103</td>
</tr>
<tr>
<td>Table 5</td>
<td>104</td>
</tr>
<tr>
<td>APPENDIX B: Classroom Observation Scale (annotated)</td>
<td>108</td>
</tr>
<tr>
<td>APPENDIX C: SEALS II Scouting Report Form</td>
<td>113</td>
</tr>
<tr>
<td>APPENDIX D: SEALS Modules</td>
<td>114</td>
</tr>
</tbody>
</table>
LIST OF TABLES

TABLE 1
Mean Total Score per Teacher Participant .................................................................43

TABLE 2
Mean Ratings of Teacher Engagement Practices for Beginning Instruction ..................44

TABLE 3
Mean Ratings of Teacher Engagement Practices During Instruction .............................46

TABLE 4
Tau-U Effect Size Analyses .............................................................................................47

TABLE 5
Change in Performance from Baseline to Intervention based on Means and Visual ..........48
LIST OF FIGURES

FIGURE 1
Engagement Practices-Total Score ................................................................. 49

FIGURE 2
Beginning-Instruction: Alignment of Instruction with Class Schedule .................. 50

FIGURE 3
Beginning-Instruction: Statement of Clear Behavioral Expectations prior to Beginning a New Task ................................................................. 51

FIGURE 4
Beginning-Instruction: Materials and Supplies Ready at the Start of Class .............. 52

FIGURE 5
Beginning-Instruction: Teacher has an Opening/Readiness Activity ...................... 53

FIGURE 6
During-Instruction: Restatement of Behavioral Expectations to Re-establish Order ........ 54

FIGURE 7
During-Instruction: Teacher Maintains High Time-On-Task Among Students .......... 55

FIGURE 8
During-Instruction: Alignment of Instruction with Class Schedule .............. 56
ACKNOWLEDGEMENT

I am grateful for the network of colleagues, friends, and family that have given their support over the last four years. From the start, I have been encouraged, challenged, and supported. It has been an honor completing this dissertation with all of you.

First, and foremost, I am deeply thankful for the mentoring that I have received. Learning from and collaborating with my advisor, Dr. David Lee, continues to be an honor and privilege. To my dissertation committee including Drs. Mary Catherine Scheeler, Paul Riccomini, and Lisa Kopp, I extend my most heartfelt thank you for your time, feedback, and encouragement. Furthermore, I am indebted to Mr. Greg Beardsley, Dr. Jean Faieta, and Dr. Phillip Belfiore for their leadership, teaching, and inspiration.

In addition, I would like to recognize with gratitude, the support of my family. A special thank you to my parents, John and Melody, for making this possible. A special thank you to my children: Jack, Hope, Halle, Chase, and Carolin for their enduring love and motivation.

Furthermore, I thank my friends for always being there, my colleagues for their understanding and empathy, and my students for helping me to be the teacher that I am. This dissertation is published with recognition of a truly remarkable team of people whom I am proud to call my colleagues, friends and family. It was made possible by Grant R305A160398 from the Institute of Education Sciences, and the opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.
INTRODUCTION

Our knowledge of science, technology, engineering, and mathematics (STEM) content is what provides us with a basis for making informed decisions regarding our personal safety, communications, and activities. (Verdugo, Navas, Gomez, & Schalock, 2012). When students are engaged in STEM learning, they increase their opportunities to develop understanding and interest in pursuing more advanced academic pursuits (e.g., algebra, statistics, geometry, chemistry, physics, calculus) particularly in the area of mathematics (Wang & Degol, 2013). Mathematics is often referred to as a “universal language” foundational to student success in the areas of science, engineering, and technology (Brower, Grimsley, & Newberry, 2007; Cozad & Riccomini, 2016). Therefore, interest and motivation in advanced mathematics topics is a significant factor influencing pursuit of post-secondary learning and careers in STEM fields (Sithole, Chiyaka, McCarthy, Mupinga, Bucklein, & Kibrige, 2017). Given the influence of mathematics knowledge on STEM learning, it is imperative that we understand the factors that influence student understanding of mathematics.

One major contributing factor to student learning of mathematics content is the classroom teacher (Darling-Hammond, 1998). Effective mathematics teachers often produce positive student outcomes. However, less effective teachers can have detrimental effects on students, particularly when students have ineffective teachers over multiple school years (Darling-Hammond, 2000). In a study examining mathematics teacher influence on student academic achievement over a period of three school years, Sanders and Rivers (1996) report a difference of fifty percentile points in student achievement when students are assigned to a sequence of less effective teachers. Understanding the characteristics of effective mathematics teachers holds great promise for improving STEM outcomes.
The National Research Council (2011) describes teachers with a “high capacity to teach in their discipline” as critical to student success in STEM education. Clearly, teacher knowledge of their content area is a critical aspect of student success, but that is only part of the story. In addition to mathematics content knowledge, other teacher characteristics have been found to influence student learning including pedagogical content knowledge (PCK), beliefs in teaching and learning, and interpretation/implementation of curricular materials (Remillard, 2005).

In the field of mathematics, the National Mathematics Council for Teachers (NCTM, 2009) uses the term Mathematical Knowledge of Teachers (MKT) in lieu of PCK. Both terms refer to teachers’ blend of content and pedagogy in a manner that enables student access to content and instruction (Shulman, 1987). Beliefs in teaching and learning also influence student outcomes in mathematics and consequently STEM education (Stipek, Givvin, Salmon, MacGyvers, 2001). Stipek et al. (2001) found a substantial correlation between teachers’ traditional beliefs and practices to teaching mathematics and decreased use of inquiry-oriented approaches promoted by the common core math standards and STEM initiatives. This leads to a potential disconnect between teachers’ interpretation and implementation of mathematics curricular materials adhering to the Common Core Mathematics Standards (CCMS; McGee, Wang, & Polly, 2013).

Common Core Math Standards (CCMS) and Professional Development

Adherence to the CCMS has, like STEM-focused initiatives, created increased attention to the need for professional development of mathematics teachers particularly in middle schools (McGee et al., 2013). First, the CCMS created a shift in the scope and sequence of content across grade levels. This shift prompted changes in curricula and/or curricular materials. Second, new content standards resulted in a strong push towards standards-based pedagogies (The
Changes in curriculum, materials, and need for effective pedagogies explains our current urgency for professional development in mathematics. This urgency is compounded in middle school as this period of time is critical to student access and advancement in higher-level mathematics courses (MET II, 2012). Failure to effectively address middle school math teachers’ professional development can have lasting detrimental effects on current STEM initiatives.

**Engaging Middle School Math Teachers in Professional Development**

On the positive side, 89% of current middle school math teachers do take advantage of math professional development (Banilower, Smith, Weiss, Malzahn, Campbell, & Weiss., 2013). Professional development that targets key skills and dispositions can have a sustainable impact on both teacher and student outcomes (Gersten, Chard, & Baker, 2000). McGee et al., (2013) and also addresses the need for professional development to engage teachers in mathematics. Engaged math teachers discuss student learning, identify common misconceptions, explicitly make connections for how mathematics concepts relate and build on each other, and develop their use of representations, tools, and strategies (McGee et al., 2013). Engaged math teachers lead to engaged students.

When teachers are not engaged, they are less attuned to student needs which can be detrimental to student outcomes and teacher goals (Hamm, Farmer, Dadisman, Gravelle, & Murray, 2011). Teacher attunement is a critical factor in identifying and breaking down barriers to learning mathematics, particularly for students with disabilities. A math teacher’s failure to provide adequate accommodations to support students with disabilities can inhibit students’ access/interaction with content and instruction. In turn, this can adversely affect students’ (a) understanding of basic concepts, (b) use of vocabulary/terminology, (c) time needed for
processing, (d) engagement in instruction, (e) activation of prior knowledge, and (f) motivation/interest (Hodge & Riccomini, 2006; Hwang & Riccomini, 2016; Ysseldyke, Thill, Pohl, & Bolt, 2005). A lack of commitment to addressing these barriers can impact students’ scores on high-stakes tests and ultimately students’ interest and participation in higher level STEM coursework, college majors, and related employment interests (Attard et al., 2013). These concerns and challenges have contributed to the allocation of resources towards teachers’ professional development in STEM areas including mathematics.

Focus and Delivery of Professional Development

Developing teachers’ capacity for effective mathematics instruction requires continued learning and professional growth through professional development programming (National Council of Teachers of Mathematics [NCTM], 1991; Scher & O’Reilly, 2007). Approaches to professional development often vary in form, features, and purpose, though ultimately the goal of professional development is to improve student outcomes. Scher and O’Reilly’s (2007) review of the literature on STEM professional development reflects a research-focus that addresses teachers’ content knowledge through various formats (e.g. workshops, coursework) paired with coaching. More recent studies reflect the need to address teachers’ pedagogical content knowledge primarily in the form of workshops paired with collaborative support and feedback (e.g., peer models, technology-enhancements, discussions, shared problem-solving) Regardless of the form of professional development, DeSimone and Parmar’s (2006) determined the three greatest areas of consideration when designing professional development include (1) how to increase in-class supports, (2) how to address immediate teacher concerns, and (3) how to use standards-based curriculum and pedagogy. Of these three considerations, Gersten et al. (2000)
note the importance of professional development in addressing teachers’ day-to-day classroom concerns as most critical.

The need for ongoing professional development for teachers is rarely contested; it is both a requirement and expectation that reflects the life-long learning philosophies encouraged in teacher preparation programs. However, there is debate as to what teachers need to know and how that knowledge should be delivered (Hill, 2007). What teachers need typically reflects four basic domains: (a) contextual knowledge - understanding of who, where, and what is taught, (b) content knowledge - understanding of subject matter, (c) pedagogical knowledge - strategies of classroom management/organization, and (d) pedagogical content knowledge - blend of content and pedagogy. Though content knowledge is frequently prioritized for STEM teachers, very few studies have examined the impact of content knowledge on student outcomes unless pedagogy is also addressed (Scher & O’Reilly, 2007). With limited research on effective professional development practices for STEM teachers, Desimone (2009) called for high quality studies that examined teachers’ development of pedagogical content knowledge. While research specific to the professional development of STEM teachers continues to reflect a lack of experimental studies, emphasis does appear to have shifted from content knowledge only to an examination of teachers’ pedagogical content knowledge. Additionally, Garet, Porter, Desimone, Birman, and Yoon (2001) has proposed five critical elements for effective professional development for STEM teachers. These include: content focus, active learning, coherence, duration, and collective participation (Garet et al., 2001; Desimone, 2009). Incorporating these elements as a base for professional development, the Supporting Early Adolescent Learning and Social Success (SEALS, Farmer, et al., 2013) model has demonstrated evidence to support the growth and development of teachers’ pedagogical content knowledge across a variety of content
Supporting Early Adolescent Learning and Social Success (SEALS & SEALS II)

What is SEALS? Supporting Early Adolescent Learning and Social Success (SEALS) is a professional development model based on research in developmental and behavioral science. SEALS training activities include pre-intervention observations and interviews with school professionals, professional development workshops, online training modules, and team and individual level implementation meetings (Farmer et al., 2013; Motoca et al., 2014). SEALS content is delivered via intervention specialist-teacher consultation. Directed consultation (DC) relies on active learning of content and collaborative participation of teachers and intervention specialists in order to affect positive changes in classrooms. Directed consultation is embedded within the SEALS training activities via directed consultation meetings (DCMs).

The purpose of SEALS is to provide teachers with skills needed to help students, particularly those in need of supports, in academic, social, and behavioral domains of functioning present in every classroom. While the skills covered in SEALS can be applied universally to enhance the learning of all students, training modules and DCMs are designed to increase teachers’ effective practice of universal, secondary, and tertiary supports across individual, classroom, and school-wide contexts in a manner that recognizes the interactional effects within and across these systems (Farmer, Farmer, Estell, & Hutchins, 2007; Farmer, Wike, Alexander, Rodkin, & Mehtaji 2015; Farmer et al., 2016).

SEALS is based on three related theories of intervention. The person-environment fit theory recognizes that an individual’s behavior is jointly determined by personal characteristics and the current environment (Eccles et al., 1993; Farmer et al., 2013). When these two factors
align, expected behaviors generally occur; however, when there is a mismatch of these factors, challenges and conflicts can occur in school environments. Practices that account for individual characteristics and environmental factors are more likely to result in preferred outcomes (Wentzel & Wigfield, 2007). The theoretical foundation of SEALS also reflects the concept of correlated constraints. Correlated constraints are described as how the presence of risk and/or protective factors inter-relate across academic, behavioral, and social functioning (Cairns & Cairns, 1994; Farmer & Farmer, 2001). The implication for practice involves a holistic approach to problem-solving recognizing the constraints of multiple factors influencing a behavior (Farmer et al., 2013). Ecological intervention posits that adjustment problems reflect problems in the interaction between the child and the ecology (Hobbs, 1982). Bronfenbrenner’s (1979) theory of ecology further addresses the dynamic changes occurring within and across sub-systems that influence development. Collectively, these theories signify how issues of academic, behavioral, and social adjustment are not static but dynamic, with bi-directional influence, across multiple sub-systems or environments.

SEALS and Multi-Tiered Systems of Support for Students. Essentially, SEALS is a means for professional development that complements multi-tiered systems of support (MTSS). As a prevention framework, MTSS allows for early identification and timely intervention for students at-risk for school adjustment problems (Hayes & Lillenstein, 2015). Hayes and Lillenstein (2015) describe the importance of aligning standards, MTSS, and teacher effectiveness systems (i.e. teacher evaluation) for improving instructional quality and ultimately student learning outcomes. However, alignment of these three initiatives is futile without also addressing effective professional development practices. Professional development opportunities
provide the means for teacher growth and change. Without change to instruction, there is less opportunity for student change in learning.

**Components of SEALS.** As a professional development model for promoting student learning outcomes, the SEALS program begins with classroom observation and interviews conducted by intervention specialists in order to determine teacher strengths and possible concerns relating to classroom social dynamics, behavior management, and academic engagement. This initial “scouting report” observation helps to identify malleable factors as well as strengths and weaknesses of teachers serving as the basis on which to build future intervention efforts (Farmer, Lee, Brooks, Chen, Moates, & Hamm, 2016). After these initial observations, teachers complete a training on the foundations of the SEALS program. The four foundational modules include (1) the theoretical underpinnings of SEALS, (2) social dynamics management, (3) competence enhancement behavior management, and (4) academic engagement enhancement.

The Theoretical Foundations module introduces teachers to the underpinnings of SEALS (person-environmental fit hypothesis, system of correlated constraints) and how these theories relate to the ecology of classrooms. The Social Dynamics Management module provides information regarding the operation of social systems within classrooms and schools. Interventions from this module work to build teacher attunement, and understanding of contingencies, social function of behavior as well as social synchrony (Hamm et al., 2010, 2014). In the Competence Enhancement Behavior Management (CEBM) module teachers learn basic classroom management strategies (e.g., Premack principle, proximity control, and high-p sequence) that can be used to support appropriate student behaviors (Conroy, Asmus, Ladwig, Sellers, & Valcante., 2004; Geiger, 2006; Lee, Belfiore, Scheeler, Hua, & Smith, 2004). The
final foundational module, Academic Engagement Enhancement (AAE), provides teachers with strategies to help focus student attention to increase academic productivity, and includes reviewing key concepts, pre-teaching, ending with success, increasing opportunities to respond, and behavioral momentum (Archer & Hughes, 2011; Sutherland, Alder, & Gunter, 2003).

These foundational modules are presented online and consist of voice over slides and activities. Each module is approximately 15 minutes in duration. After the online foundational modules are completed, the intervention specialist meets with teachers to develop an individualized plan to address classroom supports needed using SEALS supplemental modules. These brief online modules are 10-15 minutes in duration and address difficulties often encountered in classrooms (e.g., Promoting Positive Peer Cultures; Social Dynamics Management: Practical; Social Dynamics of Instructional Engagement; Routines and Schedules; Rules and Consequences; Understanding Academic and Behavioral Supporting Academic Engagement of Struggling Students; Intervening with Behavior Problems in Struggling Students; and Working with Parents).

Finally, intervention specialists meet with teachers and review the content of each completed module. The directed consultation meetings (DCMs) provide time for discussions on how to implement various features of the training within each classroom and include formative feedback from the intervention specialist. Directed consultation meetings are uniquely positioned to offer a “just-in-time” approach to professional development as students move from less (universal) to more intensive (tertiary) supports that frequently require training and adaptation (Farmer et al., 2016).

**Previous SEALS Research.** Previous research on the components of SEALS (observations/interviews, workshops, web-based modules, and directed consultation meetings)
suggest SEALS may serve as a promising approach to teacher professional development and improvement of student outcomes. The SEALS model was originally conceptualized as a refined version of the Rural Early Adolescent Learning project (Project REAL). Project REAL included a randomized control trial design involving 165 students. When compared to control schools, results indicated significant improvements in student achievement, sustained dispositions on schooling and perception of the school social context, from beginning of the year (Hamm et al., 2010). Furthermore, students in intervention schools were perceived by sixth grade teachers as friendlier and academically engaged (p < .10). In regards to teacher outcomes, greater self-efficacy and understanding of social dynamics were reported by teachers in intervention schools than in control schools when controlling for teacher age, gender, ethnicity, and fall efficacy scores (Hamm et al., 2008). Results from Project REAL support the use of the SEALS model for improving the academic, behavioral, and social adjustment of students in rural settings transitioning to middle school (Hamm et al., 2011). Building from these findings, a clustered randomized trial (CRT) was conducted involving 14 schools (7 intervention, 7 control) and 144 teachers. The CRT allowed for examination of the impact of the SEALS model in metropolitan areas. Initial findings indicated that teachers in SEALS classrooms showed a greater improvement in supportive classroom structures, behavior management, positive academic feedback than teachers in control schools. Furthermore, SEALS teachers reported greater self-efficacy with regards for meeting the needs of struggling students (Motoca et al., 2014). SEALS II, an extension of SEALS, employed a series of single-case research designs. SEALS II moved beyond universal supports to provide individualized supports (i.e., Tier 2 – Selective) through a more structured initial observation (i.e., scouting report), as well as directed consultation sessions targeted for specific students. Initial analysis of these data is currently underway. From the
SEALS II data set, an initial pilot study was completed examining the effects of SEALS on three seventh grade teachers’ use of engagement practices in inclusive STEM classrooms. Using a multiple baseline design, three math and science teachers were sequentially introduced to the SEALS intervention. A positive increase to criterion for use of engagement practices was observed across participants though mixed results were found when examining the specific engagement practice of proximity management. Overall, this study showed promise in proactively addressing academic engagement difficulties in STEM classrooms.

**Summary and Purpose of Current Study**

Improving teacher engagement practices is important as academically engaged students are commonly characterized as having higher achievement, healthier relationships with teachers and peers, and experience fewer classroom problem behaviors/disciplinary actions (Bourgeois & Boberg, 2016; Finn & Rock, 1997). As measured by the National Education Longitudinal Study (1988) data, academic engagement has a significant influence on eight grade student achievement in math and science (Singh, Granville, & Dika, 2002). A lack of student engagement in mathematics is particularly problematic as underachievement in mathematics is likely to continue and worsen over time (Montague, Krawee, Enders, & Dietz, 2014). As noted by Rizzo and Belfiore (2014) learners of mathematics must acquire and apply a wide-set of constructs across multiple domains; this is likely to impact multiple areas of STEM learning. Furthermore, acquisition and application of new skills is dependent upon mastery of previous skills (Rizzo & Belfiore, 2014). For these reasons, professional development opportunities that target teachers’ use of academic engagement practices can be used to support all students while simultaneously targeting struggling learners. To make this a reality, more research is needed on how to effectively provide professional development that increases the use and application of
evidence-based practices (EBPs) for engaging learners in math and science instruction (Lotter, Smiley, Thompson, & Dickenson 2016).

The purpose of this study is to replicate and extend previous findings that examine the effects of SEALS on teachers use of academic engagement enhancement practices. SEALS is expected to improve teacher engagement practices in inclusive, eighth grade math classrooms. To that end, the purpose of this study is to examine the impact of SEALS on the engagement practices of eighth grade math teachers with inclusive classrooms. To address this purpose, the following questions were posed:

(1) Is there a functional relation between SEALS and teachers’ overall engagement practices in classrooms?

(2) Is there a functional relation between SEALS and teachers use of engagement practices as they prepare lessons (i.e., beginning instruction)?

(3) Is there a functional relation between SEALS and teachers’ use of evidence-based engagement practices during instruction?

METHOD

Participants and Setting

Teachers of seventh and eighth grade, inclusive, STEM classrooms were invited to participate from three schools involved with project SEALS. For the purpose of consistency across teacher/classroom characteristics, participation was further narrowed to teachers of grade eight mathematics with inclusive classrooms. Six teachers, who met these criteria participated in the study (Amy, Donna, Jack, Kim, Wyatt, and Chase). The schools were located in metropolitan areas within the North East region of the United States serving grades 6-8. Each school was comparable in terms of student numbers (range = 700-900), teacher numbers (range = 50-60),
class size per teacher (range = 11-15), and number of students proficient on high stakes testing (range = 65%-80%) (“Greatschools.org,” 2018; NCES, 2015). Across schools, all teachers were certified in math and each teacher participant had three or more years’ experience. All SEALS sessions were conducted within the classroom setting. Observations occurred during morning classes while intervention sessions (i.e., directed consultation meetings) occurred during teacher preparation periods in the afternoon.

Procedures

Pre-baseline. Prior to any data collection, the intervention specialist observed each inclusive math classroom for a period of 45 minutes. The purpose of this initial observation was to identify teacher strengths and weaknesses as they related to supporting student academic engagement. This scouting report was then used to develop each teachers’ SEALS development plan (see Appendix C for the form).

Baseline. During baseline sessions, each math teacher continued with his/her typical instructional delivery (i.e. lecture, note-taking, class discussion, reviews, lab work) during their respective class. Three trained observers collected data on teachers’ instructional behaviors (see section below on dependent variables and data collection) during a continuous 30-minute observation session. No feedback or other contact occurred with teachers and intervention specialists during baseline.

Intervention. The intervention began with a brief individual orientation meeting with each teacher. During this meeting, the intervention specialist demonstrated the SEALS web site and provided the teacher with a user name and password to access training. The teacher was then asked to complete the four online foundational SEALS modules (7th & 8th Grade Theoretical Foundations; Social Dynamics Management: Theory; Competence Enhancement Behavior
Management; and Academic Engagement Enhancement) over the next week prior to the next meeting with the intervention specialist. The total duration of the four core modules was 35 minutes.

Directed consultation meetings (DCMs) began after a given teacher had completed the foundational modules. These meetings occurred using a 1:1 (face-to-face) format, every 1 to 2 weeks with the exception of periods when the school calendar had an extended break (i.e. Winter and Spring break). Meetings were typically 25-30 minutes in duration and consisted of discussions regarding teacher questions/comments about recently completed modules/activities, identification of concerns and leverage points for intervention, and ideas to best implement evidence-based strategies, given the teacher’s skill level and classroom student characteristics.

This process continued throughout the intervention phase for each teacher (See Appendix D for list of SEALS modules). Teachers selected two supplemental modules, based on classroom needs, in addition to the core modules. Selected modules were “Neurodevelopment,” “Promoting Positive Peer Cultures,” and “Supporting Academic Engagement of Struggling Learners.” Each teacher completed the module titled “Neurodevelopment” which was sixteen minutes in duration. This module conveys how adolescent brain development factors into students’ school adjustment and how teachers can implement strategies responsive to adolescents’ developmental needs (Farmer, Gatzke-Kopp, Lee, Dawes, & Talbott, 2016). Each teacher, except Chase, also selected the module “Promoting Positive Peer Cultures which was 10 minutes in duration. This module describes how teachers can influence peer cultures and strategies for creating peer cultures of effort and achievement within the classroom and across school environments. Chase selected “Supporting Academic Engagement of Struggling Students” as his second supplemental module which was fourteen minutes in duration. This module extends from the core academic
engagement enhancement module providing more extensive detail on practices to promote engagement of struggling learners in particular. These strategies include planned alternative activities, review of key concepts, pre-teaching, ending with success, opportunities to respond, the Premack principle, and behavioral momentum.

Total intervention time (including 1 hour of module completion and DCMs) was an average of four hours per participant (range 3-5). Wyatt spent 5 hours in intervention. Amy and Jack both received 4 hours of intervention. Donna and Kim each received 3.5 hours of intervention, and Chase received 3 hours of intervention.

Directed consultation meetings were conducted by graduate-level intervention specialists with teacher certification in general/special education and several years of classroom teaching experience. These intervention specialists had knowledge of developmental and ecological systems perspectives relating to the promotion of healthy school adjustment and background in evidence-based practices for students with emotional/behavioral disorders.

**Dependent Variables and Data Collection**

The measures for this study include seven behavioral indicators of academic engagement from the SEALS Observational Scale (SOS; Hamm et al., 2014; See Appendix B for Measures). This tool was developed for observation of teachers and parallels the Classroom Assessment Scoring System ([CLASS] La Paro, Pianta, & Stuhlman, 2004; Pelham, Waschbusch, & Massetti, 2008). The rating procedure follows the format of established and psychometrically validated classroom observation protocols (Weiss, Pasley, Smith, Banilower, & Heck, 2003). The engagement practices, or seven behavioral indicators, include beginning- and during-instruction activities.

Beginning-instruction practices include:
(a) teachers’ alignment of instruction with the class schedule/planning board
(b) teachers’ statement of clear behavioral expectations prior to beginning a new activity,
(c) having materials ready at the start of class, and
(d) teachers’ use of an opening/readiness activity.

During-instruction practices included:

(a) teachers’ restatement of behavioral expectations to re-establish order
(b) teachers’ maintenance of students’ time-on-task, and
(c) teachers’ frequency of review/summarizing.

The engagement practices were rated using a Likert scale (1-4) according to the following dimensions: (1) indicator is absent but necessary, (2) indicator is present but teacher orientation is negative (e.g. teacher uses proximity control [indicator is present] but the application is inappropriate or ineffective), (3) indicator is present and used favorably but inconsistently, (4) teacher behavior directly reinforces indicator (present, positive, consistent).

The annotated SOS, found in Appendix B, provides examples for each rating across all seven practices. According to Hamm et al. (2014), Cronbach’s coefficient alpha for the total scale was .92; interrater reliability (Kappa coefficient) was .88. Cronbach’s coefficient alpha for the subscale used in this study was .85.

Three trained observers, naïve to the purpose of the study and conditions, collected data throughout baseline and intervention. Two of these individuals held doctoral degrees and one individual was a graduate student. All data collectors were trained using a three-step process. First, data collectors were given thorough explanations and verbal examples for each item on the SOS. Second, the data collectors practiced data collection using sample videos of teachers instructing students in classrooms. Each data collection form was then compared to a standard,
developed by the researchers, and discrepancies were discussed. Finally, data collectors completed a competency assessment using a similar classroom video. Agreement was defined as being within one point of the standard from a scale of 1-4. All data collectors met the 90% criterion of agreement between their data sheets and the standard developed by the researchers.

Data were collected during baseline and intervention phases for each participant on average of once per week. During these sessions, data collectors observed instruction for 30-minutes and took notes. Data collectors then completed each item of the SOS immediately following the 30-minute session.

**Design**

This study was part of a larger project examining the efficacy of directed consultation, across 24 teachers using a series of single-case multiple baseline designs. Introduction of the SEALS intervention was established a priori across teachers. Teachers were randomly assigned to begin intervention prior to the start of the study. The staggering of baseline data points allowed for sequential introduction of the intervention. The sequential introduction of intervention, when using the multiple baseline design, establishes experimental control by noting (1) the change in the dependent measure as each participant moves from baseline phase to intervention phase, and (2) the change in the dependent measure for each participant once the intervention phase is introduced, while the behavior remains unchanged for other participants continuing under the baseline phase (Barlow & Hersen, 1984).

**Data Analysis**

Data for single-case design experiments are generally analyzed using visual analysis (Kennedy, 2005) and effect size calculations for single-case designs. For the purpose of this study, visual analysis included trend, level, and variability, overlap of baseline and intervention
data points, immediacy of effect, across baseline and intervention conditions (Lane & Gast, 2013).

To further support the visual analysis of each participant’s graphed data, effect size calculations were completed as recommended by Horner, Carr, Halle, McGee, Odom, & Wolery (2005). Tau-U was calculated to garner an effect size. Though several methods for calculating effect size are available for single-case designs, Tau-U is unique in its attempt to account for positive baseline trend (Parker, Vannest, Davis, & Sauber, 2011) and control for outliers in baseline that may over or under-estimate effects. Tau-U is considered a conservative effect size, relative to other options (Parker et al., 2011). According to Parker and Vannest (2009), effect size interpretations should be large effects of .91-1.0, medium effects of .66-.92, and small effects of 0-.65.

**Inter-observer Agreement**

Inter-observer agreement (IOA) was calculated by comparing the ratings of two independent observers, and then determining if the data were within 1 point of each other, for each of the seven engagement practices. IOA data were collected for 5/6 teachers during 33% of baseline sessions and 33% of intervention sessions for each participant with the exception of Chase due to scheduling conflicts. Across teachers, the median baseline agreement was 75% (range = 50%-100%). The median intervention agreement was 88% (range = 50%-100%). A correlation was also calculated to determine the relationship between observers’ scores. As previously noted, IOA data were adjusted to reflect agreement if ratings were within 1 point of each other. These percentage of agreement data are supported by a correlation of .92.

**Intervention Fidelity**
Teacher notes, intervention logs, and intervention specialist logs were used as indicators of intervention fidelity. For the modules, procedural components that were monitored included: date/time of completion, name of colleagues who may have participated in viewing the module(s), and activity/question completion. For the intervention logs, procedural components that were monitored included: date/time of completion, and feedback on the perceived effectiveness of the intervention. These logs were compared to the intervention specialist’s observation notes in order to verify completion of the intervention. The intervention specialist’s directed consultation notes were then used to confirm completion of all modules across participants.

RESULTS

Visual analysis, within and across subjects, was completed to evaluate the relation between SEALS and engagement procedures practiced by grade eight math teachers with inclusive classrooms (see Table 5). To address the first research question, analyses were performed to evaluate teachers’ engagement practices in relation to the intervention by examining the total score for all seven engagement practices. Then, analyses were completed to determine the impact of SEALS on beginning-instruction practices (alignment of instruction with class schedule; statements of clear behavioral expectations prior to start of an activity; having materials ready at the start of class; having an opening/readiness activity) followed by during-instruction practices (restatement of behavioral expectations; maintaining high time-on-task; frequent review/summarizing). Following visual analysis, Tau-U was calculated to determine effect sizes for the total score, beginning-instruction practices, during-instruction practices, and each individual engagement practice (see Table 4).

Total Score
A total score for the seven engagement practices was given using the sum of the ratings for each practice per observation session. Reaching criterion on one engagement practice is equal to a rating of “4.” Thus, the highest possible score is 28 (range 0-28). In this study, intervention was introduced sequentially (see Figure 1). Amy and Donna were introduced to intervention first. Amy’s mean total score during baseline was 24.25 (range = 23-25). This score maintained with no change in range (mean score 24.63). Donna’s baseline mean of 20.50 (range = 10-25) increased to 24.50 (range = 16-28) during intervention. Visual analysis revealed a slight accelerating trend for Amy during baseline which continued during intervention. The data from Donna reveal stable responding during the final three baseline points followed by an increase in total score across time during intervention.

Jack, Kim, and Wyatt were introduced to intervention next. Jack’s mean total score during baseline was 14.20 (range = 9-17). This score increased to 17 (range = 10-23) during intervention. However, visual analysis revealed a positive trend during baseline that continued during intervention. Kim’s baseline mean of 16.40 (range = 14-18) increased to 22.63 (range=20-26) during intervention. This difference is supported through visual analysis, which shows a change in level of performance from baseline through intervention. Wyatt’s baseline mean of 23.20 (range = 19-25) decreased to 19.75 (range = 4-25). Wyatt’s performance was variable, especially in intervention, and revealed little difference between baseline and intervention conditions.

Chase was introduced to intervention last. Chase’s mean score during baseline was 26 (range 24-28). This score increased to 27 (range 26-28) during intervention. Visual analysis reveals no difference between conditions.
In support of the visual analysis, an overall Tau-U effect size of 0.3327, indicated a modest effect of the intervention (see Table 4 for individual participant Tau-U).

**Beginning-Instruction activities**

Engagement practices for beginning instruction were present during baseline and continued with mixed results during intervention (see Table 2). From baseline to intervention, the majority of teachers maintained ratings of 3 or higher. A Tau-U score of 0.0674 indicates a negligible effect on teachers’ engagement practices for beginning instruction (see Table 4).

**Alignment of Instruction.** Again, Amy and Donna were given intervention first (see Figure 2). Amy had a mean rating of 4 (range = 3-4) during baseline that decreased to 3 (range = 1-4) during intervention. Donna’s baseline mean was also 4 (range = 3-4) and also decreased to 3 (range 1-4) during intervention. Visual analysis reveals increased variability for both Amy and Donna during intervention, with no difference in levels between baseline and intervention.

Next, Jack, Kim, and Wyatt were introduced to SEALS (see Figure 2). Jack had a mean rating of 1 (range = 1-3). This increased to 2 (range = 1-4). Kim’s baseline mean was 1 (range = 1-2). This increased to 3 (range = 1-4) during intervention. Wyatt’s baseline mean of 3 (range = 1-4) increased to 4 (range = 3-4) during intervention. Visual analysis revealed a strong intervention effect for Kim only.

Finally, Chase was introduced to SEALS (see Figure 2). Chase had a mean rating of 4 (range = 3-4) during baseline. This maintained at a mean of 4 with no range in the data during intervention. Visual analysis revealed a descending trend during the last two points of baseline. When intervention was introduced, Chase’s alignment of instruction to the class schedule increased. That said, Chase was relatively consistent at maintaining the class schedule during early baseline sessions.
**Statement of Clear Behavioral Expectations.** Teachers’ use of clear behavioral expectations was the second of four beginning instruction activities examined in this study. Amy’s mean rating during baseline was 3 (range = 3-4). This maintained with no change in range during intervention. Donna’s baseline mean was 3 (range = 3-4). This rating maintained at 3 (range =2-4) during intervention. Visual analysis reveals little change from baseline to intervention for either participant (see Figure 3).

Jack, Kim, and Wyatt received intervention following Amy and Donna (see Figure 3). Jack’s mean rating during baseline was 2 (range = 1-2) and maintained at 2 (range = 1-3) during intervention. Kim’s baseline mean was 3 (no range) which decreased to a mean of 2 (range = 1-4). Wyatt’s baseline mean was 3 (range = 3-4) which maintained during intervention with the same range in data. Visual analysis reveals more variability in performance with some increased use of behavioral expectations relative to baseline. Kim’s data reveals no trend during baseline followed by a descending trend during intervention. The data for Wyatt reveals an ascending trend in baseline, which continues during intervention.

Chase was introduced to intervention following Jack, Kim, and Wyatt (see Figure 3). Chase’s mean rating during baseline was 4 (range = 3-4). This maintained at a rating of 4 during intervention with no range in the data. Visual analysis reveals no trend during baseline or intervention phases.

**Materials and Supplies Ready at the Start of Class.** Amy and Donna both had a mean rating during baseline of 4 (range = 3-4). For Amy, a rating of 4 was maintained during intervention Donna also maintained at a rating of 4 (range = 3-4) during intervention. Visual analysis reflects little difference across phases for either participant (see Figure 4).
SEALS was introduced next to Jack, Kim, and Wyatt (see Figure 4). Jack’s mean rating during baseline was 3 (range = 2-4). This maintained with no change in range during intervention. Kim’s baseline mean was 4 (range = 3-4) which also maintained with no change in range during intervention. Similarly, Wyatt maintained a mean rating of 4 from baseline to intervention. Visual analysis reveals an ascending baseline trend for Jack followed by a slightly descending trend during intervention. For Kim, no trend is present during baseline or intervention phases. Similarly, Wyatt’s data is absent of trend across phases.

SEALS was introduced last to Chase (see Figure 4). Chase’s mean rating during baseline was 4 (no range). This rating maintained at 4 during intervention. Visual analysis reveals an absence of trend across phases for Chase.

**Opening/Readiness Activity.** The fourth and final engagement practice for beginning instruction examined teachers’ use of an opening/readiness activity at the start of class. Amy’s baseline mean rating was 4 (range 3-4). This maintained at a mean of 4 (range = 3-4) during intervention. Donna’s mean rating was the same, 4 (range 3-4) during baseline. This also maintained at 4 (range = 3-4) during intervention. Visual analysis reflects both teachers’ general use of an opening/readiness activity across baseline and intervention phases (see Figure 5).

As previously stated, Jack, Kim, and Wyatt began SEALS after Amy and Donna (see Figure 5). Jack’s mean rating during baseline was 3 (range = 2-4), which increased to 4 (range = 3-4) during intervention. Kim’s baseline mean was 4 (range = 3-4) and maintained at 4 (range = 3-4) during intervention. Similar to Kim, Wyatt’s baseline mean of 4 (no range) was maintained during intervention. Visual analysis reveals an ascending trend for Jack during baseline which continues during intervention. Kim and Wyatt’s data indicated consistent use of opening/readiness activities across baseline and intervention.
Chase began SEALS last (see Figure 5). Chase’s mean rating during baseline was 4 (no range). His mean rating maintained at 4 (no range) during intervention. Visual analysis reveals an absence of trend across phases.

**During-Instruction Activities**

Following the implementation of SEALS, teachers’ engagement practices, during instruction, maintained or improved across participants (see Table 3). A Tau-U score of 0.2689 indicates SEALS had a small effect on the teachers’ engagement practices during instruction (see Table 4).

**Restatement of Behavioral Expectations.** Amy and Donna received intervention first (see Figure 6). Amy’s mean rating was 3 (range = 3-4) during baseline and maintained at 3 (range 3-4) during intervention. Donna’s baseline mean was 3 (no range). This increased to a mean rating of 3.5 (range = 2-4) during intervention. Visual analysis reveals a stable baseline for Amy followed by an increase with the final baseline point, which maintained through much of the intervention phase. The data for Donna reveals an absence of trend during baseline followed by an ascending trend and change in level during intervention.

Jack, Kim, and Wyatt were introduced to intervention next (see Figure 6). Jack’s mean rating during baseline was 1 (range = 1-3), and Kim’s baseline mean was 1 (range = 1-2). Both Kim and Jack’s rating for restatement of behavioral expectations increased to 2 (range = 2-3) during intervention. Wyatt’s baseline mean rating of 4 (range = 3-4) decreased to a mean of 3 (range = 3-4) during intervention. Visual analysis reveals a little difference across baseline and intervention conditions for Jack. The data for Kim indicates a change in performance from baseline to intervention. Wyatt’s data were variable and showed little difference between phases.
Chase was introduced to intervention after Jack, Kim, and Wyatt (see Figure 6). Chase’s mean rating during baseline was 4 (range = 2-4). This rating of 4 maintained (range = 3-4) during intervention. Visual analysis indicates variability in the use of behavioral expectations to establish order both before and after training.

**Maintenance of Time-on-Task.** Amy’s mean rating during was baseline was 3. This increased to 4 (range = 3-4) during intervention. Kim’s baseline mean was also 3. This also increased to 4 (range = 3-4) during intervention. Visual analysis revealed an increase in use of strategies to maintain engagement of students, relative to baseline for both Amy and Donna (see Figure 7).

Following implementation with Amy and Donna, intervention was introduced to Jack, Kim, and Wyatt as previously stated (see Figure 7). Jack’s mean rating was 2 (range = 1-4). This increased to 3 (range = 2-3) during intervention. Kim and Wyatt both had a baseline mean of 3 (range = 2-4). This maintained, for Kim, at 3 (range = 2-4) during intervention. Wyatt’s rating for maintaining time-on-task increased to 4 (range = 3-4) during intervention. Visual analysis reveals an ascending trend with variability present during baseline for Jack, with no change in level during intervention. For Kim, the data revealed a descending trend during baseline followed by an ascending trend later in intervention. Wyatt’s performance was similar in level, but showed reduced variability during intervention.

Chase was the final participant to be introduced to intervention (see Figure 7). Chase’s mean rating was 3 (range = 2-4) during baseline. This increased to 4 (range = 3-4) during intervention. Visual analysis revealed similar performance late in baseline through intervention.
**Frequent Review/Summarizing.** Amy’s mean rating was 4 during baseline. A rating of 4 was maintained (range = 3-4) during intervention. Donna’s baseline mean was 3. This increased to 4 (range = 2-4) during intervention. Visual analysis revealed consistently higher performance for Amy from baseline through intervention, indicating that she reviews/summarized appropriately. Donna’s use of reviews/summaries increased after intervention was introduced (see Figure 8).

Following Amy and Donna, SEALS was introduced to Jack, Kim, and Wyatt (see Figure 8). Jack’s mean rating was 2 (range = 1-4) during baseline. This increased to 3 (range 1-4) during intervention. Kim’s baseline mean was 3 (range = 1-4). This increased to 4 (range = 3-4) during intervention. Wyatt’s mean rating was 4 (range = 3-4) during baseline. This maintained at a rating of 4 during intervention. Visual analysis reveals an ascending trend for Jack during baseline followed moderate levels of performance during intervention – indicating little change across conditions. Kim’s data showed high levels of variability in baseline, followed by consistent performance to criterion (i.e., score of 4) during intervention. Wyatt’s data were consistently high for baseline and intervention.

Chase was introduced to intervention following Jack, Kim, and Wyatt (see Figure 8). Chase’s mean rating of 4 (range = 3-4) during baseline maintained at a rating of 4 during intervention. Similarly, Chase’s data were consistently high during intervention.

**Discussion**

The purpose of this study was to evaluate Supporting Early Adolescent Learning and Social Success (SEALS) in 8th grade inclusive math classrooms. SEALS is a professional development program that provides information on evidence-based strategies to manage behavior, manage social dynamics, and increase student engagement. The program is delivered
via web-based modules along with directed consultation that helps link content to individual
teacher strengths and classroom needs. The research questions examined the effects of SEALS
on teachers’ overall, beginning-instruction, and during-instruction use of evidence-based
engagement practices.

In this study, a multiple-baseline design across participants was used to evaluate the
relation between SEALS and teacher engagement practices. Visual analysis of the data indicated
mixed, but favorable, results for certain teachers with generally small effects reported from Tau-
U calculations. Three measures were used to evaluate the effects of SEALS across participants.
First, an overall score was used to examine the general effects of SEALS on teacher use of
engagement strategies. Second, a rating of the effects of SEALS on teacher (a) alignment of
instruction with class schedule, (b) statement of behavioral expectations before starting a task,
(c) having materials ready, and (d) having an opening/readiness activity were used to examine
the effects of SEALS on engagement practices for beginning instruction. The third, and final
measure, was a rating of the effects of SEALS on teacher engagement practices during
instruction including (a) restatement of behavioral expectations, maintenance of time-on-task,
and frequent review/summarizing. The following sections address the current findings in relation
to previous SEALS research and other reform-based professional development practices as well
as the limitations and implications for research and practice.

**Intervention Effects**

**Total score.** The total score reflected overall changes in observed engagement practices
of eighth grade math teachers. Visual analysis along with Tau-U indicated small effects for Amy,
and Donna, as well as medium effects for Kim. In examining the data from this group of
inservice teachers, relatively high scores (i.e., above 25) were common across 4 of 6 participants.
in baseline, so it is not surprising that the majority of participants maintained previous practices or improved consistency of application during intervention. Two teachers in particular, Amy and Chase, used high levels of engagement strategies throughout classes and received 80% or more of the total score possible in baseline. Conversely, Kim was less consistent in implementation of evidence-based strategies during baseline. The SEALS training resulted in an increase of her total score by 25% during intervention.

**Beginning-Instruction Practices.** Similar to the total score, teacher use of beginning instruction practices was quite variable and likely based on beginning skill levels and needs within the classroom. For example, Kim rarely aligned instruction with the class schedule during baseline but increased in alignment after SEALS was implemented. On the other hand, Chase generally linked instruction to the class schedule across baseline and intervention, thus showing small effects, but he had little need for that component of the intervention. It is not surprising to see variability for aligning instruction with the class schedule as teachers either did (rating of 4) or did not (rating of 1) perform this task. Also, some teachers took a longer time to engage in appropriate beginning-instruction practices after intervention, suggesting that when intervention is needed, it may take time for training to show tangibly in the classroom. For example, only one teacher, Chase, demonstrated a mean rating at criterion (rating of 4) during baseline compared to the three teachers at criterion during the final week of intervention (Donna, Wyatt, and Chase). Similarly, in examining availability of supplies and use of opening tasks, only Jack failed to consistently display these skills during baseline. After receiving intervention, Jack’s consistency increased, although did not maintain at the levels demonstrated by other participants. While major changes in engagement practices for beginning instruction were not observed, due to teachers already at or near criterion during baseline, there are small gains made by individual
teachers primarily reflecting a change in consistency of application following the implementation of SEALS.

**During-Instruction Practices.** Slightly greater effects of SEALS were observed based on changes in teachers’ ratings for engagement practices during instruction, possibly indicating greater need for these skills across teachers. For example, Kim showed little to no evidence of “restating behavioral expectations” until after SEALS had been implemented and showed a marked improvement in reviewing/summarizing frequently during instruction.

**Previous Research on SEALS and Relationship with Present Work**

While these intervention effects appear to complement previous studies using SEALS and directed consultation, this study also includes several features that distinguish it from previous research. First, research on SEALS and directed consultation initially occurred in rural settings; however, teachers in this study were all from metropolitan locations. Second, the earlier SEALS models focused on universal intervention strategies within a MTSS format. The current study moved beyond universal supports and examined strategies for both universal and secondary-level support strategies. Too, earlier SEALS models were built on a training program that included nearly a calendar year of training. The present training occurred within the context of a 15-week program. Finally, SEALS and directed consultation were not developed to specifically address the professional development needs of teachers within a particular content area. However, this study examined SEALS in relation to the specific professional development needs of eighth grade math teachers.

While these factors are highly significant and enrich our understanding of the strengths and possible boundaries of SEALS, a major differentiating factor is the research design used in the current study. All previous research on SEALS applied group designs (e.g., RCT or CRT).
Group designs, while able to evaluate overall effects and identify possible moderating/mediating variables, may be less able to explore and evaluate intervention effects within smaller contexts, such as individual classrooms.

In this study a single-case research design was applied with the classroom teacher as the unit of analysis. Use of a multiple-baseline design across participants allowed for analysis of both the progression and post-result of intervention (Horner et al., 2005). As noted by the visual analysis of this data set, subtle changes in the quality and consistency of responding can be observed over time. For example, Donna’s restatement of behavioral expectations reflected moderate level of performance during baseline. Following the implementation of SEALS, variability increased during the first half of intervention, but became stable later in the intervention phase. As noted by Kennedy, Hirsch, Rodgers, Bruce, and Lloyd (2017), it is not unlikely for stronger outcomes to occur towards the end of a professional development opportunity as earlier training (in this case module completion and DCMs) should impact later gains. Furthermore, this design allows for varied questions than those answered in the previous group design studies (Horner et al., 2005). Examining professional development using multiple research designs can help to strengthen the evidence-base on contextual factors influencing the effectiveness of traditional and reform-based practices.

Reform-based professional development

While traditional, in-service models of professional development remain in practice (i.e. workshop, course, conference), reformed approaches to professional development have gained in momentum as evidenced by the increase in literature on teacher collaboratives (i.e. study groups), mentoring programs, resource centers, and coaching practices (Scher & O’Reilly, 2007). Like these reform-based approaches to professional development, SEALS also takes a notable
difference from traditional delivery of professional development. The most salient difference between traditional and reform-based practices is the duration of training. Current evidence suggests that the longer the duration of professional development the better the maintenance and generalization of research to practice (Kennedy et al., 2017). In this study, SEALS was implemented over a fifteen-week period.

In addition to the dimension of duration, evidence suggests four other critical components of STEM professional development including: content focus, active learning, collective participation, and coherence. While each of these dimensions appears to influence desired teacher outcomes, the major significance may be from their interaction. With regard to SEALS, each of these dimensions are present. The fluid process of SEALS permits flexibility for content-specific professional learning that is active, collaborative, and coherent.

Identified critical elements of STEM professional development can be applied via many formats. Pending the purpose and function, these formats may or may not be efficient and effective. Studies examining which contextual factors may warrant specific professional development practices is an area that has historically relied on expert recommendation or theoretical understanding. While multiple forms of professional development serve varied purposes, SEALS is unique in that application takes into account other forms of ongoing professional development allowing for connections from the PD learning environment(s) to the classroom with feedback and ongoing support. For example, SEALS incorporated school district initiatives and teacher-set goals used for evaluation purposes, as well as school-wide behavioral supports, and other initiatives. The coordinated effort lead to teacher-driven, targeted professional development.
SEALS is also unique from other professional development practices in that it is a holistic approach. Research as well as practice of professional development typically targets teachers’ application of an evidence-based practice or fidelity of an evidence-based practice. SEALS is much more comprehensive allowing for professional development to be individualized and capitalize on teacher’s strengths as well as areas of high-leverage within and across school environments. By not taking a deficit approach, or only identifying a singular approach to combat one teacher “weaknesses,” teachers’ conceptualization of professional development begins to shift from product to process. The advantage of this enlightened conceptualization of teacher professional development is teacher ownership of their professional learning. SEALS provides a structure that facilitates, encourages, and scaffolds self-directed learning of teachers.

**Limitations and Future Research**

The potential of SEALS as a reform-based approach to professional development warrants review of the limitations of this study and considerations for future research. Limitations are discussed in terms of the research design, participants, and sensitivity of measures. In regard to the research design, a multiple baseline design shows replicable effects across participants; however, experimental control can be a concern when intervention is not sequentially introduced. The difference of only one data point following introduction of the intervention across legs of the design makes it more difficult to confirm experimental control particularly when trend is present during baseline. A second concern is the amount of overlapping data from baseline to intervention across participants. This limited the opportunity to show large, quantifiable differences as many teachers were already at or near criterion prior to the start of intervention (i.e., ceiling effects). Relatedly, the teachers in this study were assigned to inclusive classrooms and may have had more advanced skills or expertise for teaching in
diverse classroom settings. This may explain why so many of the participants were close to mastery during baseline. A fourth concern is that the sensitivity of the measure may have impacted our ability to detect subtle changes to teachers’ engagement practices. The SEALS Observational System was designed to evaluate each behavioral indicator within the overall context of a given classroom. It is possible that examination of other variables from the SOS may show different results and more closely align with areas needing improvement. Relatedly, the nature of the data made it difficult for maintaining high levels of observer agreement and no IOA data was collected for the last participant, Chase. This measure did perform well in prior group designs but may not have been as suitable for use in single-case designs where direct counts of behavior (e.g., the frequency of use of a given strategy) are often used. A fourth concern relates to documentation of intervention implementation. Teacher module completion was collected through self-report and verified within the intervention specialists’ log notes. This may have limited the accuracy of recorded time that elapsed between actual module completion and start of directed consultation meetings. Also, no exact protocol for directed consultation meetings was stipulated which may have influenced the range of outcomes and effects across dependent measures.

While the purpose of this study was to examine teacher behavior, following studies should help to establish the link between teacher and student behavior. Also, the use of structured intervention log note may assist with future documentation of DCMs for fidelity purposes. Furthermore, the focus of this study specifically examined the effects of SEALS on eighth grade math teachers’ engagement practices. Examining the impact of SEALS across STEM domains may provide further insights on how to best support the growing need to fill and retain STEM teachers with respect to the contextual challenges of teaching in these domains.
Implications for Practice

Given the shortage of teachers with advanced knowledge of mathematics concepts in middle school settings, prioritizing professional development is a necessity for meeting current STEM initiatives. When teachers, or school districts, are determining participation in a professional development opportunity, attention and priority is often placed on getting evidence-based practices into the mainstream use of teachers.

Addressing the “what” of professional development is indeed crucial; however, other factors, including how professional development is delivered, require equal attention. Results of this study, in combination with previous research on SEALS, raises three major considerations regarding selection and participation in professional development opportunities of middle school math teachers. The first consideration is evaluating whether current or future forms of professional development adhere to the critical components of STEM professional development. The second consideration is whether current or future professional development opportunities take a holistic approach or focus more directly on one particular practice or strategy. A third consideration is how the professional development experience may influence teachers’ beliefs and conceptualization of professional development. In other words, will the professional development practice enhance or diminish a learning-oriented approach of teachers with regard to their own learning (i.e. product-vs-process). Involvement in professional development practices that (1) reflect the current evidence-base, (2) take a holistic account of variables impacting classroom instruction, and (3) explicitly prioritize the process of professional learning over the product are likely to impact teacher self-efficacy (Hamm et al., 2014; Motoca, et al., 2014) and ultimately retention efforts.

Conclusion
Taken overall, the data from this study highlight the individualized nature of how pedagogical skills play out in classrooms and that matching professional development to each teacher’s needs is critical. Results of the present study indicate SEALS produced change in the quality and consistency of eighth grade math teachers’ engagement practices in inclusive settings for teachers who needed those specific skills. When teachers effectively apply practices that increase student engagement, changes in academic, behavioral, and social outcomes are more likely to reflect improvement. Student improvement is the ultimate goal of teacher professional development. This study contributes to the growing evidence-base of professional development and our understanding of professional development practices for middle school math teachers. Future consideration of SEALS as an approach to support middle school math teachers reflects a commitment to understanding and establishing teacher learning opportunities that embrace our current understanding of effective professional development for teachers.
References


scouting report process. *Intervention in School and Clinic, 52*(2), 67-76.
doi:10.1177/1053451216636073


doi:10.1016/j.beth.2010.08.006


Table 1

Mean Total Score per Teacher Participant

<table>
<thead>
<tr>
<th></th>
<th>Baseline (range)</th>
<th>Intervention (range)</th>
<th>Change</th>
<th>Final Week Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amy</td>
<td>24 (23-25)</td>
<td>24 (23-25)</td>
<td></td>
<td>25</td>
</tr>
<tr>
<td>Donna</td>
<td>21 (10-25)</td>
<td>25 (16-28)</td>
<td>+</td>
<td>28</td>
</tr>
<tr>
<td>Jack</td>
<td>14 (9-17)</td>
<td>17 (10-23)</td>
<td>+</td>
<td>21</td>
</tr>
<tr>
<td>Kim</td>
<td>16 (14-18)</td>
<td>23 (20-26)</td>
<td>+</td>
<td>21</td>
</tr>
<tr>
<td>Wyatt</td>
<td>23 (19-25)</td>
<td>20 (4-25)</td>
<td>-</td>
<td>25</td>
</tr>
<tr>
<td>Chase</td>
<td>26 (24-28)</td>
<td>27 (26-28)</td>
<td>+</td>
<td>27</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>20 (9-28)</strong></td>
<td><strong>23 (4-28)</strong></td>
<td><strong>+</strong></td>
<td><strong>25</strong></td>
</tr>
</tbody>
</table>
Table 2

Mean Ratings of Teacher Engagement Practices for Beginning Instruction

<table>
<thead>
<tr>
<th></th>
<th>Baseline (range)</th>
<th>Intervention (range)</th>
<th>Change</th>
<th>Final Week Score</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Alignment of instruction</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amy</td>
<td>4 (3-4)</td>
<td>3 (1-4)</td>
<td>-</td>
<td>3</td>
</tr>
<tr>
<td>Donna</td>
<td>4 (3-4)</td>
<td>3 (1-4)</td>
<td>-</td>
<td>4</td>
</tr>
<tr>
<td>Jack</td>
<td>1 (1-3)</td>
<td>2 (1-4)</td>
<td>+</td>
<td>4</td>
</tr>
<tr>
<td>Kim</td>
<td>1 (1-2)</td>
<td>3 (1-4)</td>
<td>+</td>
<td>4</td>
</tr>
<tr>
<td>Wyatt</td>
<td>2 (1-3)</td>
<td>1 (0-2)</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>Chase</td>
<td>3 (1-4)</td>
<td>4 (3-4)</td>
<td>+</td>
<td>3</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td><strong>Statement of behavioral expectations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amy</td>
<td>3 (3-4)</td>
<td>3 (3-4)</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Donna</td>
<td>3 (3-4)</td>
<td>3 (2-4)</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Jack</td>
<td>2 (1-2)</td>
<td>2 (1-3)</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Kim</td>
<td>3</td>
<td>2 (1-4)</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>Wyatt</td>
<td>3 (3-4)</td>
<td>3 (3-4)</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Chase</td>
<td>4 (3-4)</td>
<td>4</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td><strong>Materials ready at start of class</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amy</td>
<td>4 (3-4)</td>
<td>4 (3-4)</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Donna</td>
<td>4</td>
<td>4 (3-4)</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Jack</td>
<td>3 (2-4)</td>
<td>3 (2-4)</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Kim</td>
<td>4 (3-4)</td>
<td>4 (3-4)</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Wyatt</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Chase</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>4 (3-4)</td>
<td>4(3-4)</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>----------</td>
<td>---------</td>
<td>--------</td>
<td>-----</td>
<td>---</td>
</tr>
<tr>
<td>Amy</td>
<td>4 (3-4)</td>
<td>4 (3-4)</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Donna</td>
<td>4 (3-4)</td>
<td>4 (3-4)</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Jack</td>
<td>3 (2-4)</td>
<td>4 (3-4)</td>
<td>+</td>
<td>4</td>
</tr>
<tr>
<td>Kim</td>
<td>4 (3-4)</td>
<td>4 (3-4)</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Wyatt</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Chase</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td><em>Average</em></td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>0</td>
</tr>
</tbody>
</table>
Table 3

*Mean Ratings of Teacher Engagement Practices during Instruction*

<table>
<thead>
<tr>
<th></th>
<th>Baseline (range)</th>
<th>Intervention (range)</th>
<th>Change</th>
<th>Final Week Score</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Restatement of behavioral expectations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amy</td>
<td>3 (3-4)</td>
<td>3 (3-4)</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Donna</td>
<td>3</td>
<td>4 (2-4)</td>
<td>+</td>
<td>4</td>
</tr>
<tr>
<td>Jack</td>
<td>2 (1-3)</td>
<td>2 (2-3)</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Kim</td>
<td>1 (1-2)</td>
<td>2 (2-3)</td>
<td>+</td>
<td>2</td>
</tr>
<tr>
<td>Wyatt</td>
<td>4 (3-4)</td>
<td>3 (3-4)</td>
<td>-</td>
<td>4</td>
</tr>
<tr>
<td>Chase</td>
<td>4 (2-4)</td>
<td>4 (3-4)</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Maintains high TOT among students</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amy</td>
<td>3</td>
<td>4 (3-4)</td>
<td>+</td>
<td>4</td>
</tr>
<tr>
<td>Donna</td>
<td>3</td>
<td>4 (3-4)</td>
<td>+</td>
<td>4</td>
</tr>
<tr>
<td>Jack</td>
<td>2 (1-4)</td>
<td>3 (2-3)</td>
<td>+</td>
<td>3</td>
</tr>
<tr>
<td>Kim</td>
<td>3 (2-4)</td>
<td>3 (2-4)</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Wyatt</td>
<td>3 (2-4)</td>
<td>4 (3-4)</td>
<td>+</td>
<td>4</td>
</tr>
<tr>
<td>Chase</td>
<td>3 (2-4)</td>
<td>4 (3-4)</td>
<td>+</td>
<td>4</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td>3</td>
<td>4</td>
<td>+</td>
<td>4</td>
</tr>
<tr>
<td><strong>Frequent review/summarizing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amy</td>
<td>4</td>
<td>4 (3-4)</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Donna</td>
<td>3</td>
<td>4 (2-4)</td>
<td>+</td>
<td>4</td>
</tr>
<tr>
<td>Jack</td>
<td>2 (1-4)</td>
<td>3 (1-4)</td>
<td>+</td>
<td>3</td>
</tr>
<tr>
<td>Kim</td>
<td>3 (1-4)</td>
<td>4 (3-4)</td>
<td>+</td>
<td>4</td>
</tr>
<tr>
<td>Wyatt</td>
<td>4 (3-4)</td>
<td>4</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Chase</td>
<td>4 (3-4)</td>
<td>4</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td>3</td>
<td>4</td>
<td>+</td>
<td>4</td>
</tr>
</tbody>
</table>
Table 4

*Tau-U Effect Size Analyses*

<table>
<thead>
<tr>
<th></th>
<th>Amy</th>
<th>Donna</th>
<th>Jack</th>
<th>Kim</th>
<th>Wyatt</th>
<th>Chase</th>
<th>CUM.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Overall Score</strong></td>
<td>0.1563</td>
<td>0.5313</td>
<td>0.4286</td>
<td>1</td>
<td>-0.5250</td>
<td>0.4444*</td>
<td>0.3327*</td>
</tr>
<tr>
<td><strong>Beginning Instruction</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.0674</td>
</tr>
<tr>
<td>Alignment</td>
<td>-0.1563</td>
<td>-0.4167</td>
<td>0.0476</td>
<td>0.8438*</td>
<td>-0.4000</td>
<td>0.1481</td>
<td>0.0470</td>
</tr>
<tr>
<td>Clear Statement of Behavioral Expectations</td>
<td>-0.1250</td>
<td>-0.0417</td>
<td>0.3214</td>
<td>-0.5000</td>
<td>0.1714</td>
<td>0.1111</td>
<td>0.0038</td>
</tr>
<tr>
<td>Materials Ready at the Start of Class</td>
<td>0.1250</td>
<td>-0.2500</td>
<td>0.4667</td>
<td>0.0750</td>
<td>0</td>
<td>0</td>
<td>0.0713</td>
</tr>
<tr>
<td>Has an Opening activity</td>
<td>0.2500</td>
<td>0</td>
<td>0.600</td>
<td>0.2500</td>
<td>0</td>
<td>0</td>
<td>0.1823</td>
</tr>
<tr>
<td><strong>During Instruction</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.2689*</td>
</tr>
<tr>
<td>Restatement of Behavioral Expectations</td>
<td>0.2500</td>
<td>0.5000</td>
<td>0.3000</td>
<td>0.8500*</td>
<td>0.8500*</td>
<td>0.0370</td>
<td>0.4949*</td>
</tr>
<tr>
<td>Maintenance of TOT</td>
<td>0.7500*</td>
<td>0.7500*</td>
<td>0.1429</td>
<td>0.2000</td>
<td>0.300</td>
<td>0.0741</td>
<td>0.3673*</td>
</tr>
<tr>
<td>Frequent Review/Summarizing</td>
<td>-0.2500</td>
<td>0.7500**</td>
<td>0.2857</td>
<td>0.5000</td>
<td>0.2000</td>
<td>0.2222</td>
<td>0.2772**</td>
</tr>
</tbody>
</table>

*Note.* CUM. = combined data from all four participants. * = significance at the .05 level. ** = significance at the .1 level.
Table 5

*Change in Performance from baseline to intervention based on means and visual analysis*

<table>
<thead>
<tr>
<th>Total Score</th>
<th>Alignment</th>
<th>Behavioral Expectations Stated</th>
<th>Materials Ready</th>
<th>Opening Activity</th>
<th>Restatement of Behavioral Expectations</th>
<th>Maintains Time on Task</th>
<th>Frequent Review</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amy</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Donna</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Jack</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Kim</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Wyatt</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Chase</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

*Note.* /+/ = improvement, /-/ = maintenance current performance,
Figure 1. Engagement Practices – Total Score
Figure 2: Beginning-Instruction: Alignment of Instruction with Class Schedule
Figure 3. Beginning-Instruction: Statement of Clear Behavioral Expectations prior to beginning a New Task
Figure 4. Beginning-Instruction: Materials and Supplies Ready at the Start of Class

Amy

Donna

Jack

Kim

Wyatt

Chase
Figure 5: Beginning-Instruction: Teacher has an Opening/Readiness Activity
Figure 6. During-Instruction: Restatement of Behavioral Expectations to Re-establish Order

Baseline | Intervention

Amy
Donna
Jack
Kim
Wyatt

Chase

Sessions
0 1 2 3 4 5 6 7 8 9 10 11 12 13
Rating for Restatement of Behavioral Expectations
Figure 7. During-Instruction: Teacher Maintains High Time-On-Task Among Students

Baseline

Intervention

Amy

Donna

Jack

Kim

Wyatt

Chase

Sessions

Rating for Maintains High Time-on-Task

0 1 2 3 4 5 6 7 8 9 10 11 12 13
Figure 8. During-Instruction: Teacher Frequently Reviews/Summarizes
Appendix A

Literature Review

Professional Development for STEM Teachers: A Review of the Literature

Karen Rizzo

The Pennsylvania State University
Abstract

Improving teaching and learning is largely influenced by professional development (Darling-Hammond, 2000; Desimone, 2009). While many forms of teacher professional development exist, and offer advantages/disadvantages, the evidence to support the array of practices remains limited (Desimone, 2009; Yoon, Duncan, Lee, Scarloss, & Shapley, 2007). This is particularly concerning given the immediate need for professional development of science, technology, engineering, and mathematics (STEM) teachers (Desimone, 2009). Building from the work of Scher and O’Reilly (2007), this paper describes the characteristics and impact of professional development practices on teacher practice and student outcomes in STEM classrooms. Results indicate a prioritized focus on practices aimed to improve STEM teachers’ pedagogical content knowledge with desired teacher and student outcomes when applying five critical elements of professional development. Recommendations for practice and future research are discussed.

Keywords: professional development, science, math, consultation, STEM
The report, *A Nation-At-Risk* (Gardner, Larsen, Baker, Campbell, & Crosby, 1983), called for reform of the nation’s math and science educational programs noting our international ranking based on performance on international assessments (i.e., Trends in International Mathematics and Science Study, TIMSS; Programme for International Student Assessment, PISA). It was additionally noted that there has been a lack of progress in closing achievement gaps between various groups (e.g., gender, race/ethnicity, disability, etc.) in national science, technology, engineering, and mathematics (STEM) assessments (Department of Education, 2015a, 2015b; Hwang & Riccomini, 2016; Rizzo & Taylor, 2016). Though progress is noted by the National Assessment of Education Progress (NAEP), eighth grade science score increase in 2015 (National Center for Education Statistics [NCES], 2015), continued efforts are required to address the ongoing concerns in, and related to, STEM education.

Great concern continues over growth and sustainability of several STEM related issues including: the future workforce, science literacy of the citizenry, enrollment/graduation rates of STEM majors, the rigor of high school graduation requirements in STEM, inclusion and access to advanced STEM related courses, and overall STEM achievement (National Research Center [NRC], 2011; National Science Foundation [NSF], 2014; Vilorio, 2014). To address these concerns, new standards in math (i.e., Common Core Math [CCM] or similar) and science (i.e., Next Generation Science Standards [NGSS] or similar) were developed. As of 2017, approximately 43 states have adopted/maintained adoption of the CCM standards, and at least 12 states have adopted/maintained the NGSS or similar versions of these standards (Coburn, Hill, & Spillane, 2016). Higher standards in STEM have prompted the need for new curricula and curricular materials. Changes to STEM standards and curricula challenged traditional math and
science instructional approaches, creating a sense of urgent need for professional development in STEM. A rise in professional development (PD) initiatives prioritized STEM education for all learners.

**The Need for Professional Development in STEM**

In education, the classroom teacher influences all learning; the content areas of math and science are no different (Cohen & Hill, 2000; Darling-Hammond, 1998). This is particularly concerning in STEM classrooms for three reasons. First, the US is currently facing a shortage of STEM teachers (Ingersoll, 2003). With fewer individuals enrolling in teacher training programs, there is cause for concern over our ability to fill these positions. Second, STEM teachers lack adequate preparation for implementing the NGSS upon leaving most teacher preparation programs (Desimone, Porter, Garet, Yoon, & Birman, 2002). Math teachers report feelings of inadequacy due to the current push toward conceptual understanding that conflicts with how they learned to teach and/or new methods of inquiry (Desimone et al., 2002). Additionally, some secondary education majors lack preparation for teaching diverse student populations (Kahn & Lewis, 2014). Third, a large number of practicing teachers report attitudes and self-efficacy issues relating to their ability to use inquiry-based methods and properly meet the needs of diverse learners (Kahn & Lewis, 2014). Tackling the issues surrounding teacher shortages, insufficient preparation, and low self-efficacy of STEM teachers is a national priority. As noted by Hwang and Riccomini (2016) this is particularly true for mathematics; Understanding of mathematics can be universally applied in science, technology, and engineering fields. Professional development opportunities offer great promise for addressing each of these concerns.
While STEM PD is advantageous to overcoming the barriers linked to teaching in these areas, it is not without its own challenges. These challenges vary from state to state; however, a few common issues exist. According to Wei, Darling-Hammond, and Adamson (2010), limited induction/mentoring programs, declined intensive PD around content knowledge, and limited time/resource allocation of PD all contribute to the current issues with STEM education PD. Wei et al. (2010) further notes the decline from 2004 to 2008 in use of effective PD models. When less effective PD models are implemented (e.g. “one shot” workshops), the generalization of skills taught in PD to the actual classroom is unlikely; even if teachers try implementing what is taught during PD, fidelity can be lost (Odom, 2009).

**Effective Professional Development in STEM**

Debate continues regarding the effectiveness of various PD models/forms (i.e., how PD should happen) (Desimone, 2009). However, there is consensus on effective core and structural features of PD (i.e., what PD should include) (Garet et al., 2001; Wei et al., 2010). Additionally, research suggests that PD should focus on one, or any combination, of the four knowledge bases (i.e., contextual, content, pedagogical, and/or pedagogical content)(McGee et al., 2013).

**What STEM PD should include.** Five critical components of effective PD include: (a) content focus, (b) active learning, (c) coherence, (d) duration, and (e) collective participation (Garet et al., 2001; Desimone et al., 2002; Desimone, 2009). Given the general agreement of research and expertise on these characteristics, each critical component holds value in PD for STEM educators.

**Content focus.** Content and content pedagogy were listed by Blank, de las Alas, and Smith (2008) as one PD characteristic that impacts practice and student learning. Interestingly, Wei et al. (2007) notes that content-focused professional development yields differentiated
results based on teachers’ years of experience with evidence suggesting less impact on new teachers with less than three years’ experience. This is particularly concerning in STEM education which has already shown that teachers are not leaving preparation programs with sufficient content knowledge and pedagogy to address ever-changing adoption of new standards (Kahn & Lewis, 2014).

**Active learning.** Wei et al. (2007) concludes that PD should include active learning of both content and how to teach specific content. Active learning is conceptualized as teachers “learning by doing (inquiry)” in which math and science teachers are given opportunities to investigate and construct their knowledge (Loucks-Horsley et al., 1996). More broadly defined within the context of professional development, active learning can include engagement in, or with, inquiry, curriculum discussions, planning, and/or practice (Darling-Hammond, 1998; Loucks-Horsley et al., 1998; & Penuel & Means, 2004). Active learning opportunities in PD are greatly enhanced by the formative feedback provided (Scheeler, Ruhl, & McAfee, 2010). In STEM education, opportunities to receive feedback on active learning experiences may not be readily available for several reasons such as limited resources or less-intensive PD models. This may be more concerning in rural and urban districts which may not have well established mentoring programs which include active learning opportunities (Wei et al., 2007).

**Coherence.** Mentoring programs are one aspect of coherence, a key characteristic in PD. Coherence is described as the integration and alignment of PD with school/district initiatives (e.g. changes to curriculum, standards, or assessment practices). Coherence can also be thought of as the establishment and maintenance of an active system of support across stakeholders (e.g. teachers, administrators, community members) (Loucks-Horsley et al., 1996).
**Duration.** There is little debate that PD which is more intensive in duration yields better results in teacher practice and student achievement (Wei et al., 2007). Current recommendations suggest PD should be at least 1 year in duration; however, this general recommendation does not consider contextual factors; a broader look at duration of PD with regard to specific teacher needs may warrant consideration for various durations particularly if viewing PD as a continuum. In a meta-analysis of STEM PD studies by Scher and O’Reilly (2007), they indicate that the greatest gains in student math achievement occurred when PD occurred over multiple years. Similarly, student achievement was greatest in science when PD was at least one year duration though no statistical significance was reported for multiple years of PD in science (Scher & O’Reilly, 2007).

**Collective participation.** Central to the idea of collective participation is collaboration. Evidence suggests that opportunities to collaborate with colleagues as a significant influence on teacher retention (Borko et al., 2008; Grossman et al, 2001; Slavit et al. 2013). Teachers who have opportunities to collaborate with peers, colleagues, and related stakeholders are less likely to leave the profession (Slavit et al., 2013). Loucks-Horsley et al. (1996) characterizes collective participation in terms of professional learning communities supported by school and district administrations in which continued learning, risk-taking, reform, innovation, and a shared vision for math and science education are a constant part of the culture.

**How STEM PD should happen.** While there may be general consensus of effective core and structural features of STEM professional development, there are a wide-range of forms of professional development. Forms of professional development range from traditional approaches (e.g. workshops, courses, and conferences) to reform-based approaches (e.g. teacher study groups, teacher collaboratives, mentoring, internships, and resource centers) (Desimone et al.,
There is little support or belief in effectiveness for using an in-service model alone without follow-up coaching, consultation, or opportunities for feedback (Wei et al., 2007). Forms of professional development continue to vary though most of the literature between 1990 and 2004 started to examine the practice of coaching (Scher & O’Reilly, 2007).

**Why STEM PD should happen.** The various forms and features of STEM PD address how professional development “looks” and/or is delivered but equally important is the concentrated focus of professional development, or what is covered. Abell et al. (2007) describes at least four basic knowledge bases including: contextual, content, pedagogical, and pedagogical content knowledge. Contextual knowledge refers to teachers’ understanding of context (i.e. who, where, and what is taught). Pedagogical knowledge includes strategies of classroom management and organization that are cross-curricular (Gess-Newsome, 1999). Content knowledge refers to teachers’ understanding of the subject matter (Hill et al. 2005). Content knowledge is sometimes considered a pre-requisite to pedagogical content knowledge but Lee et al. (2007) caution against the assumption that CK leads to PCK. Since Shulman’s (1987) conceptualization of PCK, it has been considered the highest priority of teacher preparation though variations of interpretation do exist (Gess-Newsome, 1999). PCK can be described as a blend of pedagogy and content relating to (a) instructional strategies, (b) students, (c) curriculum, and (d) assessment (Grossman, 1996). In other words, it is what makes content accessible to students (Shulman, 1987). Since the National Science Foundation’s (2005) call for research on STEM PD for enhancing PCK, the concept and term have been much more widely applied particularly in terms of science education. However, the National Council for Teachers of Math (NCTM, 2009) does not make a distinction between CP and PCK. Instead, research regarding the preparation of math teachers may use the term mathematical knowledge for teachers (MKT) which combines the terms of CK
and PCK. Drawing from the research examining math teachers’ knowledge and impact on practice and student outcomes, two distinctions have been suggested. While CK does influence higher student achievement, PCK has shown to be more effective in obtaining desired student achievement outcomes. Furthermore, only PCK has evidence that suggests desired effects on teacher practice (Hill, Rowan, & Ball, 2005; Baumert et al., 2010). Simply stated, changes to these knowledge bases are thought to influence teacher practice and thus produce changes in student learning (Baumert et al., 2010; Scher & O’Reilly, 2007).

Improving the learning and engagement of diverse student groups is a major factor influencing both generalization of evidence-based practice and reduced teacher attrition (Gersten Chard, & Baker, 2000.). Teachers instructional behaviors are typically reinforced when students experience success thus new teaching practices are more likely to be sustained overtime. When new research-based practices are generalized by the classroom teacher, a positive chain of events is set in motion with regard to teacher retention efforts, sustained best practice, and continual student (and teacher) learning. These points are illustrated and elaborated by Desimone (2009), Scher & O’Reilly (2007), and Loucks-Horsley et al. (1996) in their respective work toward a common vision and conceptual framework for professional development of STEM teachers.

The common vision and conceptual framework for PD of STEM teachers is certainly a positive direction for improving teacher learning. Unfortunately, the research on teacher learning and professional development of STEM teachers is sparse (Scher & O’Reilly, 2007). This contributes to our continued reliance on past practice and expert recommendations. While many individuals accept the necessity for professional development, the inherent value of PD (related to teacher knowledge, practice, and student outcomes) requires thoughtful examination of best practice particularly with regard to the current call for effective STEM teachers. Responsive to
this need, Scher and O’Reilly’s (2007) evaluation of the evidence-base on STEM PD included studies from 1990 through 2004. Their review yields three relevant and important findings for moving forward. First, the intensity of PD makes a difference in student achievement in STEM courses. Second, better student outcomes occur when teachers’ PCK is addressed versus CK alone. Third, no determination is confidently suggested in regards to the form of STEM PD due to limited data and lack of study rigor. However, most of the studies did include a coaching/consultation component (Scher & O’Reilly, 2007). Noting a limited research-base, Scher & O’Reilly (2007) express the need to continue exploration of the variables impacting STEM PD calling for further evaluation of the components, form, and focus of interventions and their impacts on mediating and long-term teacher and student learning outcomes.

**Purpose of the Review**

A decade since Scher & O’Reilly’s (2007) review of STEM professional development practices, STEM teacher and student learning continues to be a national priority (Sithole et al., 2017). Addressing the learning needs of STEM teachers and their respective students requires diligent and vigilant analysis of STEM professional development practices. Replicating and extending the work of Scher and O’Reilly (2007), this review synthesizes the findings of research studies investigating professional development for STEM teachers and their subsequent changes in practice as well as student outcomes. These results will serve to inform recommendations for professional development targeting STEM teachers working with diverse student populations.

Specific questions include:

1. What are the characteristics of research studies from 2005 to 2017 that examine STEM professional development?
2. What are the intervention characteristics of each research study that focuses on STEM professional development from 2005 to 2017?

3. What outcomes (teacher and/or student) are in STEM PD studies published from 2004 to 2017?

4. What are the effects on student and/or teacher outcomes for STEM professional development research studies published from 2005 to 2017?

5. What is the quality of experimental and quasi-experimental STEM professional development research studies?

**Method**

**Inclusion Criteria**

Included studies met the following criteria: (a) published between 2005-2017, (a) occurred in US, (c) middle school, STEM classrooms, and were (d) located in peer-reviewed, scholarly journals. Additionally, included studies investigated the effects of STEM PD programs on teacher knowledge, practice, and/or student achievement. For the purpose of this review, measures of teacher outcomes include researcher-developed assessments, survey, observation, and video data. Student outcome measures included end of year grades and achievement data. Similar to Scher and O’Reilly (2007), quasi-experimental studies with a control group were included due to limited available experimental studies meeting remaining search criteria.

**Literature Search**

Three electronic databases, via the Pennsylvania State University library system, and a hand search of selected journals relevant to professional development and STEM was completed. Search terms included: professional development, science, math, and STEM. PsychINFO, ERIC, and ProQuest databases yielded a total of 1,903 possible articles. A hand search of *Teaching and*
Teacher Education, Teacher Education and Special Education, and the Journal of Science and Math resulted in 53 possible articles. Spanning across seven journals, 11 articles by 41 different authors were identified as meeting the criteria for inclusion in this review of the literature.

Analysis of the Studies

Several analyses were included in this review of the literature. Descriptive analyses were conducted on the studies in the areas of teacher/setting characteristics, research design/quality characteristics, core/structural features, and measures of teacher and student outcomes. For studies that specifically examined teacher outcomes, analysis of the knowledge-base addressed and alignment to recommended core and structural features were analyzed. For studies that included student measures, analysis of type of student outcomes were addressed (i.e. academic, social, behavioral).

Effect Size Calculations

Effect sizes for group design studies, was calculated using Hedges’ g. Hedge’s g calculations account for the overestimation that occurs when calculating ES using studies with small sample sizes (Hedges, 1981). As suggested by Cohen (1988), effect size interpretations should be large effects of .80 and above, medium effects at .50 to .80, and small effects at .50 and below. For single case design studies, Tau-U effect size calculation was conducted. The Tau-U calculation determines the percentage of non-overlap between phases controlling for positive baseline trend (Parker et al., 2011). According to Parker and Vannest (2009), effect size interpretations should be large effects of .91-1.0, medium effects of .66-.92, and small effects of 0-.65.

Study Quality Indicators
The quality of research for group design studies, was analyzed using the quality indicators proposed by Gersten et al. (2005). Gersten et al. (2005) provides a list of essential quality indicators relating to conceptualization, participants, intervention implementation/nature of the comparison group, outcome measures, and data analysis. There is also a list of seven desirable quality indicators (Gersten et al. 2005). Gersten et al. (2005) proposes that a study of high quality includes all but one essential quality indicators and at least four desired quality indicators. Acceptable studies are thought to include all but one essential quality indicator and at least one desired quality indicator (Gersten et al., 2005).

The quality of research for single-case design studies, was analyzed using the quality indicators proposed by Horner et al. (2005). Horner et al. (2005) provides twenty-one quality indicators representing the seven areas of participants, setting, dependent variable, independent variable, baseline, experimental control/internal validity, external validity, and social validity. These quality indicators serve as the guideline for determining if the methodological rigor of a single-case design study is acceptable (Gersten et al., 2005).

**Results**

In this review, the author examined studies from 2005 to 2017, analyzing STEM PD according to participants, settings, research designs, interventions, and outcomes. Along with extracting characteristics of each study, effect size analyses and study quality was examined. Effect size analyses consisted of Hedge’s g for group design studies and Tau-U for single case design studies. Study quality was analyzed using suggestions for group designs (Gersten et al., 2005) and single case designs (Horner et al., 2005) respectively.

**Descriptive Analyses of Study Characteristics**
Across the 11 studies included in this review, there were a total of 266 teacher participants teaching mainly in middle school general education settings (Anderson & Hoffmeister, 2007; Blanchard, Prevost, Tolin, & Gutierrez, 2016; Hanegan, Friden, & Nelson, 2009; Kennedy, Hirsch, Rodgers, Bruce, & Lloyd, 2017; Lotter, Smiley, Thompson, & Dickenson, 2013; Maher, Palius, Maher, Hmleo-Silver, & Sigley, 2014; Patel, Franco, Miura, & Boyd, 2012; Penuel, McWilliams, McAuliffe, Benbow, Mably, & Hayden, 2009; Scheeler, Congdon, & Stansbery, 2010; Singer, Lotter, Feller, & Gates, 2011; Walker, Recker, Robertshaw, Sellers, & Leary, 2012). A quasi-experimental design (with comparison) group was used most commonly to examine teachers’ pedagogical content knowledge. The majority of studies reflected desired teacher and student outcomes relating to inquiry-based teacher practices in math and science. Participant information consisted of the number of participants per study and their respective content area, grade level, experience, and training/certification. Setting characteristics included information on the school (e.g. public, private, size), area (e.g. urban, rural, suburban), and classroom (e.g. inclusive, non-inclusive).

Participants. A total of 266 participants were included in the review of 11 studies. Of the seven studies that specified teacher grade level, it appears that the sample represents a fairly even distribution across middle school grades 6-8 (Blanchard et al., 2016; Hanegan et al., 2009; Lotter et al., 2013; Patel et al., 2012; Penuel, 2009; Scheeler et al., 2010; Singer et al., 2011. Teachers of math included 50% of the participant population with the remaining majority teaching science (28%) and a smaller number of teachers (22%) providing math and science instruction. Additionally, a small number (5%) of teachers from other content area were included in this review as their results could not be teased out (Anderson & Hoffmeister, 2007; Maher et al., 2014; Walker et al., 2012). While these teachers were not math or science content teachers, they
were integrating math and science content into their instruction (See Table 1). In addition, 56% of teachers had a minimum 5 years’ experience teaching.

**Setting.** The majority of the studies (10/12) reported taking place in a public school; the remaining 2 studies did not specify (Kennedy et al., 2017; Lotter et al., 2013). Five studies reported on the percentage of the students receiving free/reduced lunch (Blanchard et al., 2016; Kennedy et al., 2017; Lotter et al., 2013; Penuel, 2009; Singer et al., 2011). From these studies, 90% (4/5) included schools with a population in which 50% or more of the students were provided free/reduced lunch. Urban, rural, and suburban settings were specified in 5/11 studies (Blanchard et al., 2016; Penuel, 2009; Scheeler et al., 2010, Singer et al., 2011; Walker et al., 2012; Of the studies reporting this information, there is an equal distribution. Similarly, only half the studies specified if the intervention occurred in inclusive classrooms. Though the remaining half of studies did not specify inclusion, it is inferred that these settings included students receiving special education services or thought to be at-risk based on current inclusion rates (Heasley, 2016). Overall, professional development of math/science teachers was examined primarily in inclusive public school classrooms in which 50% of the population may have been receiving free/reduced lunch across urban, rural, and suburban settings.

**Research design.** The majority of the studies implemented a group design with the exception of Scheeler et al. (2010) which implemented a single-case design (See table 1). Of the eleven studies, four studies used experimental designs (Kennedy et al., 2017; Patel et al., 2012; Penuel, 2009; Scheeler et al., 2010), five studies were quasi-experimental with a control group (Anderson & Hoffmeister, 2007; Lotter et al., 2013; Maher et al., 2014; Singer et al., 2011; Walker et al., 2012), and two studies used a mixed methods design (Blanchard et al., 2016; Hanegan et al., 2009).
Intervention Characteristics

Analysis of intervention characteristics included the form, critical components (content focus, duration, coherence, active learning, and collective participation), and type of knowledge-base targeted. Across the twelve included studies, workshops with collaborative support and feedback, employing each of the five critical components, and targeting teachers’ pedagogical content knowledge were most common.

Form. Three basic forms of intervention were implemented across studies (coursework, workshop, or resource-based professional development). Across the included studies, 9/11 could be considered reform-based approaches as their components went beyond traditional coursework, “one-shot” in-service, or conference approach. Anderson & Hoffmeister (2007) was not considered a reform-based approach to STEM professional development as the intervention was solely coursework with no additional component (i.e. collaborative/follow-up support). Similarly, Patel et al. (2012) was not considered reform-oriented as the workshop format did not incorporate any other components beyond that of traditional formats. Of the studies implementing reform-based approaches, Hanegan et al. (2009) implemented the use of coursework with collaborative support. The majority of reform-based intervention studies (5/9) included a workshop with collaborative support (Blanchard et al., 2016; Lotter et al., 2013; Penuel, 2009; Singer et al., 2011 Walker et al., 2012). The remaining three studies implemented a resource-based approach to professional development with collaborative support (Kennedy et al., 2017; Maher et al., 2014; Scheeler et al., 2010).

Critical components. All five critical components (content focus, active learning, coherence, duration, and collective participation) were identified in each study. Across studies certain patterns and differences emerged.
**Content focus.** Content focus includes both the content and/or the content pedagogy of the professional development. In this review, 42% (4/11) of the studies focused on professional development in the area of math (Anderson & Hoffmeister, 2007; Maher et al., 2014; Patel et al., 2012; Scheeler et al., 2010) 33% (4/12) focused on science content (Hanegan et al., 2009; Lotter et al., 2013; Penuel, 2009; Singer et al., 2011) and 25% (3/12) focused on math and science (Blanchard et al., 2016; Kennedy et al., 2017; & Walker et al. 2012). All studies prioritized pedagogical content knowledge in math and/or science. Additionally, three studies included a focus on teachers’ understanding of the content in math (Anderson & Hofmeister, 2007; Patel et al. 2012) and science (Hanegan et al., 2009).

Studies only involving math teachers included content relating to operations or problem solving; professional development was not specific to a sub-category of math (e.g. geometry or algebra) nor was delineation made as to what math courses were taught by participants of these studies. Similarly, studies only including science teachers did not involve professional development content specific to one branch of science (e.g. anatomy or physics) with the exception of Penuel et al. (2009) which only involved the content area and teachers of Earth science. Studies involving both math and science teachers did not have a content-specific focus in math or science but rather technology and/or positive behavior intervention supports.

**Duration.** Duration of STEM professional development ranged widely from two hours to three years with 60% of the studies including a duration less than one year and the remaining 40% of studies with a duration of at least one year or longer (up to 3 years). Approximately half of the studies were reportedly less than a year in duration ranging from two hours to three months (Anderson & Hoffmeister, 2007; Kennedy et al., 2017; Patel et al., 2012; Scheeler et al., 2010; Singer et al., 2011; Walker et al., 2012). Of these six studies, Anderson and Hoffmeister
(2007), Scheeler et al. (2010), and Walker et al. (2012) were three months in duration. Patel et al. (2012) and Kennedy et al. (2017) were 1 week in duration (hours varied from 2 to 40), and Singer et al. (2011) was three weeks (105 hours). The remaining five studies were at least one year in duration (range of 1-3 years). Three studies were one year in duration (Hanegan et al., 2009; Lotter et al., 2013; Penuel, 2009) while one study was 3 years in duration (Blanchard et al., 2016).

**Coherence.** Each of the twelve studies included in this review acknowledged a system of coherence. The vast majority of studies include an active system of support ([e.g. university partners, community partners, teacher collaboratives] Blanchard et al., 2016; Hanegan et al., 2009; Kennedy et al., 2017; Maher et al., 2014; Patel et al., 2012; Penuel, 2009; Scheeler et al., 2010; Singer et al., 2011, Walker et al., 2012). Two studies were supported by school district initiatives (Anderson & Hoffmeister, 2007; Lotter et al., 2013). Only one study involved a co-commitment of mentoring and active system of support (Scheeler et al., 2010).

**Active learning.** Active learning was represented by teachers’ engagement in inquiry (4), curriculum (2), discussions (4), planning (2), and practice (6). Interventions that involved teachers’ engagement in inquiry allotted opportunities for problem solving (Anderson & Hoffmeister, 2007; Maher et al., 2014) or investigations (Penuel, 2009). Penuel (2009) and Patel et al. (2012) both engaged teachers with the curriculum. Both studies had teacher participants take the role of students but only Penuel (2009) also had teachers involved in investigations, using the curriculum, as teachers. Teacher engagement through discussion included dialogue regarding research (Anderson & Hoffmeister (2007) or specified content (Hanegan et al., 2016; Kennedy et al., 2017; Maher et al., 2014). While ¼ of these studies provided face-to-face discussions, Hanegan et al. (2009) made use of a listserv to facilitate ongoing dialogue.
Furthermore, three-fourths of the studies paired discussion activity with other forms of teacher engagement; Only Kennedy et al. (2017) provided active learning solely through teacher discussion. Teacher engagement in planning took the form of teams collaborating on lesson/unit plans following a teaching, or mock teaching situation (Penuel, 2009; Singer et al., 2011). Approximately half of the studies engaged teachers in practice. Practice varied according to teachers’ engagement with technology (Blanchard et al., 2016; Scheeler et al., 2010; Walker et al., 2012), or inquiry-based practices (Lotter et al., 2013; Singer et al., 2011).

**Collective participation.** Collective participation occurred in five specific structures. (1) Group problem solving, (2) technology-enhanced collaborations, (3) collaborative discussions, (4) peer models plus feedback, and (5) lesson/unit planning. Group problem solving in which teachers collaborated on solving math problems were components of two studies (Anderson & Hoffmeister, 2007; Maher et al., 2014). Maher et al. (2014) paired this with technology-enhanced collaboration. Technology-enhanced collaboration was the second form of collective participation and used by Blanchard et al. (2016) and Hanegan et al. (2009) with variation. A third structure to collective participation was collaborative discussions. Collaborative discussions were the sole structure for collective participation in Walker et al (2012). It was used in combination with TEC by Hanegan et al. (2009). Patel et al. (2012) included collaborative discussion opportunities though not specified or structured and therefore not deemed as including this component. A fourth structure found was the use of peer models plus feedback. This structure was implemented as the exclusive form of collective participation in two studies (Kennedy et al., 2017; Walker et al., 2012) and paired with other structures by Lotter et al. (2013) and Scheeler et al. (2010). Additionally, Kennedy et al.(2017) specified the form of peer modeling plus feedback as the practice of coaching. Collaborative inquiry groups
were used solely by Lotter et al. (2013) in combination with the peer model plus feedback and technology-enhanced collaborations. Lesson/unit planning in teams was the final structure of collective participation used independent of the other structures in both Penuel (2009) and Singer et al. (2011).

**Knowledge-Base.** Identified knowledge-bases include content knowledge (CK), pedagogical knowledge (PK), contextual knowledge (CTK), and pedagogical content knowledge (PCK). In this review, STEM professional development has focused solely on the knowledge-base described as pedagogical content knowledge ([PCK]see table x). As a blend of pedagogy and content, PCK includes: knowledge of content and students (KCS), knowledge of content and curriculum (KCC), knowledge of content and teaching (KCT), and knowledge of content and assessment (KCA). Of the twelve studies, almost 67% (8/12) focused on KCT. The remaining studies focused on KCC (Patel et al., 2012; Penuel, 2009) and KCS (Anderson & Hoffmeister, 2007; Maher et al., 2014). No studies in this review focused on KCA.

**STEM Professional Development Outcomes and Measures**

**Outcomes and measures.** Across the twelve studies included in this review diverse teacher outcomes were reported with variation even within the identified categories. All studies reported teacher outcomes and 5/11 report student outcomes. Reported teacher outcomes related to content knowledge (Anderson & Hoffmeister, 2007; Patel et al., 2012), knowledge of the learner (Maher et al., 2014), planning (Lotter et al., 2013; Penuel, 2009), inquiry-based instruction (Blanchard et al., 2016; Hanegan et al., 2009; Lotter et al., 2013; & Singer et al., 2011), and other instructional practices (Kennedy et al., 2017; Scheeler et al., 2011; & Walker et al., 2012). Student outcomes related to behavior (Kennedy et al., 2017; Walker et al., 2012) and academics (Blanchard et al., 2016; Hanegan et al., 2009; Penuel, 2009; & Walker et al., 2012).
**Reported Teacher Outcomes and Measures.** Each of the 11 studies examining teacher outcomes reported desired results. Teacher outcomes included content knowledge, lesson/unit planning, inquiry-based practices, and other teacher practices (i.e. positive behavior supports, use of three-term-contingencies, application of technology, or problem-based learning).

Both studies examining teachers’ content knowledge focused on building conceptual understanding through either a math methods course (Anderson & Hoffmeister, 2007) or curriculum workshop (Patel et al., 2012). Pre- and post-tests reflected statistically significant gains in mean scores for each study. Maher et al. (2014) examined teachers’ knowledge of the learner (recognition of students’ math reasoning skills) using a researcher-developed assessment. Based on pre-assessment data, growth rates of the experimental groups were significantly higher than the growth rates reported for the control groups.

Lotter et al. (2013) and Penuel (2009) examined changes in teachers’ lesson/unit planning. Using observational tools including the Reformed Teacher Observation Protocol (RTOP; Lotter et al., 2013) and the Structured Observation Protocol (SOP; Penuel, 2009) both studies reported observed changes in teacher planning though the changes in pre- and post-assessments were not statistically significant for either study.

Four studies examined teachers use of inquiry-based practices using two similar observation measures ([RTOP] Blanchard et al., 2016; Lotter et al., 2013 & Singer et al., 2011) ([COP] Hanegan et al., 2009). Hanegan et al. (2009) reported a significant increase in teachers addressing higher cognitive skills and using open-ended questions in comparison to the control group, and Singer et al. (2011) reported sustained reform-based practices. Both of these studies reported statistically significant differences in pre- and post-mean scores of inquiry practices. Blanchard et al. (2016) and Lotter et al. (2013) did not report statistically significant changes in
teachers’ use of inquiry practices. Furthermore, Blanchard et al.’s (2016) post-mean scores did not meet criteria for indicating reform-based instruction. However, improved efficiency and effective-use of technology was reported. Unlike Blanchard et al. (2016), Lotter et al. (2013) post-test mean scores did reflect criteria of reform-based instruction though no statistically significant improvement was reported.

Other instructional practices targeted for improvement included positive behavior supports (Kennedy et al., 2017), completion of three-term contingency trials (Scheeler et al., 2010), and knowledge/use of technology and/or problem-based learning (Walker et al., 2012). Kennedy et al. (2017) reported higher-use of positive behavior supports by teachers participating in the CAP-TV plus coaching condition versus the control condition. Scheeler et al. (2010) reported a percentage of completed TTC trials per observational session. Reported findings reflect consistently higher levels of completed TTC trials, maintained and generalized post-intervention across participants when bug-in-ear technology is paired with peer coaching. Based primarily on self-reported survey data, Walker et al. (2012) reports large effects for the use of both technology and technology/problem-based learning conditions on teachers’ knowledge and use of technology though stronger effects appeared in the latter condition.

**Reported Student Outcomes and Measures.** Student behavioral outcomes included observed time-on-task (Kennedy et al., 2017) and self-reported student survey data on two behavioral indicators (not specified) (Walker et al., 2012). Academic outcomes were more prevalent and measures included end-of-grade scores (Blanchard et al., 2016), an increase in student-use of open-ended questions (Hanegan et al., 2009), higher rates of interpretations from students given a model or analogy in science (Penuel, 2009), and self-reported student survey
data on 3 behavioral indicators of knowledge acquisition (not otherwise specified) (Walker et al., 2012).

**Effect Size Analyses**

Effect sizes were calculated for all group design studies using Hedge’s g, and Tau-U was calculated for single-case research designs. For group designs, studies range in effect size from .03 to 1.88 (See Table 4). Studies ranged from small effects to large effects (Cohen, 1988). Of the ten group design studies included in this, effect sizes were able to be obtained for seven studies (Blanchard et al., 2016; Hanegan et al., 2009; Kennedy et al., 2017; Maher et al., 2014; Patel et al., 2012; Singer et al., 2011; Walker et al., 2012). Effect sizes were calculated for both teacher and student reported data. Small effects were found for teacher and student outcomes of three studies (Blanchard et al., 2016; Patel et al., 2012; & Walker et al., 2012). Medium effects were found for teacher outcomes in Hanegan et al. (2009), and large effects were found for teacher and student outcomes of three studies (Kennedy et al., 2017; Maher et al., 2014, and Singer et al., 2011) The remaining three group design studies did not provide enough information to determine effect size (Anderson & Hoffmeister, 2007; Lotter et al., 2013; Penuel, 2009). Scheeler et al (2010) implemented a single-case research design; Tau-U was calculated resulting in an effect size of 1 which indicates a large effect size.

**Quality of Research Studies**

Each of twelve studies were reviewed for quality using the quality indicators developed and recommended by Gersten et al. (2005) for group research designs, and Horner et al. (2005) for single-case research designs. Gersten et al. (2005) suggests for a study to be of high quality, it must meet all but one “essential quality indicator” and at least four “desirable quality indicators.” Studies may be considered acceptable if all but one “essential quality indicators” are met and at
least one “desirable quality indicator. Based on the findings in this review 6/10 group design studies are of high quality (Hanegan et al., 2009; Kennedy et al., 2017; Lotter et al., 2013; Maher et al., 2014, Patel et al., 2012; & Walker et al. 2012) and 4/10 studies (Anderson & Hoffmeister, 2007; Blanchard et al., 2016; Penuel, 2009; & Singer et al., 2011) do not meet the suggestions by Gersten et al. (2005) to be considered of high or acceptable quality because each is missing at least two of the essential quality indicators. Scheeler et al. (2010) meets the criteria described by Horner et al. (2005) to be considered as a high quality study for single-case research.

Discussion

The purpose of this review was to examine the current research on professional development for STEM teachers by replicating and extending the work of Scher and O’Reilly (2007). Each of the twelve studies meeting inclusion criteria between 2005-2017 were descriptively analyzed according to the study, intervention, and outcome characteristics. Furthermore, effect size and quality of the research are reported. Analysis of these factors address the path of the core conceptual framework for professional development developed by Desimone (2009). Though the number of studies meeting criteria for inclusion for this review are small, results indicate primarily veteran teachers, within diverse public school settings, across the United States participated in quasi-experimental, group design studies between the period of 2005-2017. Of these eleven studies, three are both of high quality and yield large effects on teacher and student outcomes (Kennedy et al., 2017; Maher et al., 2014, & Scheeler et al., 2010). Results of this review suggest evidence in support of three forms of professional development incorporating each of the five proposed critical components when working with experienced teachers to improve PCK. The following sections puts this information within the context of the
Limitations of this review, implications for practices, and suggestions for continued research efforts are also provided.

**Study characteristics**

Analysis of study participants, setting(s), and research design was completed for each of the twelve included studies. Based on the results of this review, STEM professional development has focused primarily on math and/or science teachers with more than five years’ experience in middle school classrooms. Similar to STEM professional development studies dating back to 1990, settings continue to represent diverse student populations in urban, rural, and suburban public school settings across the United States (Scher & O’Reilly, 2007). With regard to research design, it is encouraging to have found five experimental studies as this research methodology is sparse in the literature on STEM professional development (Wayne et al., 2008; Scher & O’Reilly, 2007; Yoon et al., 2007). It is also worth noting that all five experimental studies focused on middle school teachers. Middle school represents a critical period for addressing the concerns surrounding STEM education (e.g. access to higher level STEM courses, science literacy, career interest/exploration). Given the contrast between how veteran teachers were taught math/science and shift to more inquiry-based curriculums and focus on conceptual understanding of content, it is promising, and not surprising, that research efforts have focused on these study characteristics.

**Intervention characteristics**

Intervention characteristics were descriptively analyzed in regards to the form, critical components, and knowledge-base(s) of each STEM professional development. From 1990-2004, the form of STEM professional development varied greatly across studies though components of coaching (or consultation), added to workshop-based professional development opportunities...
were more prominent than other forms of professional development (Scher & O’Reilly, 2007) Odom et al. (2009) differentiates these coaching/consultation practices from team-based discussions noting the need for a third party person with expertise. Forms of professional development including coaching, or consultation, are described as “enlightened professional development approaches (Odom et al., 2009).”. Since 2004, the dominant form of STEM professional development appears to be the workshop plus collaborative support and feedback. Given that the term “coaching” was only specifically mentioned in one study (Patel et al., 2012), it was more appropriate to code this added feature of the workshop in terms of collaborative support and feedback. The prominence of coaching/consultation prior to 2004 and collaborative support/feedback described in each of the twelve studies of this review is significant given the desired outcomes to teacher and student learning. Central to the studies reviewed here and by Scher and O’Reilly (2007) involving coaching/consultation, is the feedback provided. As reported by Scheeler, Ruhl, and McAfee (2010) immediate feedback is effective for changing teacher behaviors. Interestingly, each of the studies using a resource-based form paired with collaborative support and feedback demonstrated large effects (Kennedy et al., 2017; Maher et al., 2014; Scheeler et al., 2010).

Regardless of form, each of the included twelve studies contained most of the five critical elements of STEM professional development described by Desimone et al. (2002) and supported in the literature (Garet et al., 2001; Loucks-Horsley, 1996; Borasi & Fonzi, 2002; Wei et al., 2010) with variation. First, Content focus (content and content pedagogy) of each study was specific to the content area(s) of the participants of each respective study. Given the push for improving the content knowledge of STEM teachers, it is interesting that only 3/12 studies measure related teacher outcomes (Anderson & Hoffmeister, 2007; Patel et al., 2012; Hanegan et
al, 2009). The second critical component is duration with general acceptance that the longer the duration of professional development, the better the outcomes. As reported by Scher and O’Reilly (2007), this is particularly true for studies on math professional development. However, the studies included in this review reflect a wide range in duration and each demonstrated desired outcomes. The third critical component, coherence, signifies the importance of working within a system to create sustainable change. Coherence is reflected in these twelve studies analyzed via school district initiatives or other active systems of support (e.g. professional learning communities, grade level teams, school-university partnerships). Wei et al. (2010) promotes the use of intensive forms of professional development specifically noting the importance of mentoring yet only Scheeler et al. (2010) incorporates this practice. Active learning is the fourth critical element. While this term is often associated with learning by doing, a broader definition was used to capture the active participation of STEM teachers involved in the studies. Active learning, or participation, was primarily done through practice though engagement in inquiry and/or discussion was also prominent. The fifth and final critical element of collective participation was evident across studies in the form of group problem solving, TECs, collaborative discussions, or peer models paired with feedback. The various forms of collective participation were rarely used in isolation of one another. While each of the critical components is present across studies, the variation among them suggests that professional development can range in application across these elements.

While form addresses what STEM professional developments should look like, and the critical components reflect how it should happen, it is the knowledge-base that signifies the purpose, or why, professional development should occur. While content knowledge, contextual knowledge, and pedagogical knowledge are each uniquely important to teaching and learning,
PCK is often considered the gold standard for impacting medial and long-term outcomes for teachers and students (Hill, 2007; Kennedy, 1999; Scher & O’Reilly, 2007). The call for studies that addressed PCK appears to be answered as 10/12 studies implemented STEM professional development for this purpose. More specifically, Knowledge of Content and Teaching (KCT) served as the dominant purpose with very few studies targeting knowledge of content and students (KCS), knowledge of content and curriculum (KCC), and no studies addressing knowledge of content and assessment (KCA).

**Outcome characteristics**

Teacher outcomes were more prominently addressed as only half of the studies reported student outcomes. Reported teacher outcomes were diverse though the majority of studies focused on inquiry-based instructional practices. Student data focused on achievement and was measured using various tools (e.g. EOGs, state tests, self-report surveys, and observation tools). Only 5/12 studies report both teacher and student outcomes. Of these five studies, only one demonstrates large effects on both teacher and student outcomes.

**Effects of STEM professional development**

An almost equal ratio of small to large effect sizes were found for teacher and student outcomes across the eight studies providing enough information to calculate an effect size on group design studies. Scheeler et al. (2010) also demonstrated a large effect size for single-case research. The large effects found in Kennedy et al. (2017), Maher et al. (2014), and Scheeler et al. (2010) provides preliminary support for these interventions for STEM professional development.

**Quality of the Research**
While the quality of research was evident in the majority of studies for this review, a lack of high quality, experimental studies described by Desimone et al. (2002) and Scher and O’Reilly (2007) continues to limit our understanding of best practice for STEM professional development. Lack of reporting on the fidelity of intervention implementation appeared to be the most neglected quality indicator followed by a lack of reported effect sizes.

**Limitations**

There are several limitations of this review. First, the purpose of this review was to examine professional development practices of STEM teachers. Each of these studies employed teachers of math and science but very few teachers of technology were included and no teachers specifically teaching engineering skills/concepts were included. Furthermore, a second limitation may be the interpretation as how STEM is defined varies in the literature. A third and final limitation of this review is the variance within the independent measures. This variance coupled with the limited sample of studies meeting inclusion criteria limits interpretations of the results from this review.

**Implications for future practice and research**

While a vast array of formal and informal professional development opportunities are available and applied, Desimone (2009) suggests using a core conceptual framework for measuring critical features of professional development rather than examining types of professional development activities. This review provides support for reform-based practices showing evidence of the five generally accepted critical features of professional development.

While form will vary, and critical features will be debated, very few people contest the practice of professional development even when evidence of little to no effect on teacher practice is thought to be present (Jacob & McGovern, 2015). Instead, Jacob and McGovern (2015)
challenge those in the field of education with aligning PD investments with those that may
demonstrate improved outcomes. Their recommendation for school systems to (1) re-define how
to help teachers, (2) take inventory of current PD practices, and (3) address how to bring
improved practice to scale stems from the belief that “what” is delivered via any form of
professional development currently has limited or no evidence of maintained or generalized
effect on teacher behavior. They further report that half of the teachers sampled reported little to
no need for help in improving teaching practices.

Though limited evidence of “what works” for professional development may be present,
the research-base is wide and does provide insights for future directions. It is encouraging to see
the increase in experimental studies since Scher and O’Reilly’s (2007) meta-analysis. As
expressed by Odom et al. (2005), it is necessary for continued research to vary in methodologies
as varied research designs answer different questions that cumulatively offer an improved
understanding. This is no different, or less critical, in relation to STEM professional
development. When examining the impact of professional development and teacher practice and
student outcomes, care must be taken with regard to methodological issues that may arise
(Wayne et al., 2008).

**Conclusions**

Given the expenditures on professional development, a clear evidence-base is warranted
as to what constitutes best practice for STEM professional development (Jacob & McGovern,
2015; Wayne et al., 2008). The study characteristics reflected in this review, highlight an
increased effort to understanding STEM professional development for middle school teachers
with diverse classrooms across the United States. The large effects shown for teacher and student
outcomes, in 4/11 studies, provides further support for incorporating the five critical components
of professional development described by Desimone et al. (2002) and supported in the literature (Loucks-Horsley et al, 1996; Wei et al., 2010). Building from the current evidence-base and continued trends in STEM education, the need for high quality professional development of STEM teachers will continue to rise. As the literature-base of what works in STEM education continues to grow, equivalent must be our commitment to ensuring application and sustainability of evidence-based practices for teaching in STEM areas to all learners.
References


doi:10.3102/0013189X027001005


doi:10.1177/002221940003300505


doi:10.1016/j.beth.2010.08.006


Retrieved from http://scholarworks.rit.edu/jsesd/vol19/iss1/2


<table>
<thead>
<tr>
<th>Study</th>
<th>Intervention Characteristics</th>
<th>Dependent Variable</th>
<th>Design/Measure</th>
<th>Teacher Outcomes</th>
<th>Student Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anderson &amp; Hoffmeister (2007)</td>
<td>Math Methods Course</td>
<td>Content understanding of MSM teachers</td>
<td>Quasi-experimental; Pre- and post-test</td>
<td>Increase of conceptual knowledge</td>
<td>na</td>
</tr>
<tr>
<td>Blanchard et al. (2016)</td>
<td>Teacher technology-enhanced PD</td>
<td>Teachers beliefs/practice of guided-Inquiry investigations using hand-held probe-ware</td>
<td>Mixed methods; Observation tool</td>
<td>Significant increase in teachers’ reported comfort in using technology though no significant change in RTOP score.</td>
<td>Higher EOG scores for 8th grade students when instructed by teachers participating in intervention</td>
</tr>
<tr>
<td>Hanegan et al. 2009</td>
<td>Course +support</td>
<td>Student participation and teachers’ application of inquiry practices</td>
<td>Mixed methods study with matched control group</td>
<td>Teachers provided significantly more time for knowledge representation activities and reported more use of open-ended questioning than teachers in comparison group</td>
<td>Increase in students use of open-ended questions</td>
</tr>
<tr>
<td>Kennedy et al. 2017</td>
<td>CAP-TV plus coaching</td>
<td>PBIS (i.e. OTR, Specific Praise, Pre-correction, Student engagement,)</td>
<td>Experimental</td>
<td>Teachers who received the intervention used significantly more of the PBIS practices than</td>
<td>Students of intervention teachers were thought to be more on-task than students with comparison group</td>
</tr>
</tbody>
</table>
teachers in the comparison group | teachers though this data is not statistically significant

Table 1 (continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Intervention Method</th>
<th>Outcome Measure</th>
<th>Design</th>
<th>Intervention Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lotter et al. (2013)</td>
<td>Summer Institute with follow-up support</td>
<td>Lesson design/implementation</td>
<td>Quasi-experimental; observation tool</td>
<td>¾ intervention groups of teachers were able to carry out inquiry-based lessons; Increase in observation tool scores indicating reform-based lesson though change in score not statistically significant.</td>
</tr>
<tr>
<td>Maher et al. (2014)</td>
<td>VMC videos and discussions</td>
<td>Teacher recognition of students’ math reasoning</td>
<td>Quasi-experimental; Researcher-developed assessment tool</td>
<td>Growth rates of teachers in intervention were significantly higher than for those teachers in the comparison group</td>
</tr>
<tr>
<td>Patel et al. (2012)</td>
<td>CMC Workshop</td>
<td>Math content knowledge and attitudes</td>
<td>Experimental; Pre- and Post-test</td>
<td>Statistically significant gain from pre- to post-test</td>
</tr>
<tr>
<td>Penuel (2009)</td>
<td>IES, ESBD, and Hybrid Approach</td>
<td>Earth Science lesson plan/implementation</td>
<td>Experimental; observation tool</td>
<td>Observed differences in planning across the curriculums though no statistical significance. Students with Teachers in the ESBD or hybrid conditions were more likely to promote student</td>
</tr>
<tr>
<td>Study</td>
<td>Intervention Description</td>
<td>Frequency of completed three-term contingency trials</td>
<td>Multiple-baseline design across participants</td>
<td>Percentage of completed TTC trials increased at onset of intervention across each participant and maintained at higher level during fading, maintenance, and generalization phases</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>------------------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------</td>
<td>---------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Scheeler et al. (2010)</td>
<td>Peer coaching with bug-in-ear technology</td>
<td>Frequency of completed three-term contingency trials</td>
<td>Multiple-baseline design across participants</td>
<td>Percentage of completed TTC trials increased at onset of intervention across each participant and maintained at higher level during fading, maintenance, and generalization phases</td>
</tr>
<tr>
<td>Singer et al. (2011)</td>
<td>Summer institute with sustained classroom support</td>
<td>Inquiry-based Practice</td>
<td>Quasi-experimental</td>
<td>Statistically significant improvement in total as well as 4/5 sub-test scores on the RTOP from pre-to post-assessment</td>
</tr>
<tr>
<td>Walker et al. (2012)</td>
<td>Technology-related teacher PD (TTPD): tech only vs. tech + PBL</td>
<td>Knowledge/use of online learning resources</td>
<td>Quasi-experimental</td>
<td>Both conditions reflected pre-post gains but teachers in the tech + PBL had larger gains for self-reported knowledge and externally rated use of PBL</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Students of teachers in tech+PBL condition, reported increased gains in behavior, knowledge and attitude (tech only condition – reflected gains in attitude)</td>
</tr>
</tbody>
</table>

Table 1 (continued)
Note. PD = professional development; CAP-TV = Content Acquisition Podcasts for Teachers with Embedded Modeling Video; MSM = middle school math project; VMC = video mosaic collaborative; CMC = connecting math concepts; IES = Investigating Earth Systems; ESBD = Earth Science by Design; tech = technology; PBL = problem-based learning; PBIS = positive behavior intervention supports; STEAM = science, technology, engineering, art, and math; RTOP = Reformed Teaching Observation Protocol; TTC = three term contingency trials; EOG = end of grade; CBMs = curriculum-based measures; OTR = opportunities to respond
Table 2

*Intervention characteristics*

<table>
<thead>
<tr>
<th>Study</th>
<th>CF</th>
<th>Duration</th>
<th>Coherence</th>
<th>Active Learning</th>
<th>Collective Participation</th>
<th>Reform-Oriented</th>
<th>PCK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anderson &amp; Hoffmeister (2007)</td>
<td>M</td>
<td>&gt;yr.</td>
<td>SDI</td>
<td>Inquiry; Discussion</td>
<td>Group problem solving</td>
<td>No-Coursework only</td>
<td>KCS</td>
</tr>
<tr>
<td>Blanchard et al. (2016)</td>
<td>M</td>
<td>&lt;yr.</td>
<td>ASS</td>
<td>Practice</td>
<td>TEC</td>
<td>Workshop + Collaborative support</td>
<td>KCT</td>
</tr>
<tr>
<td>Hanegan et al. (2009)</td>
<td>S</td>
<td>=1 yr.</td>
<td>ASS</td>
<td>Inquiry; Discussion</td>
<td>Collaborative discussions &amp; TEC</td>
<td>Coursework + CS</td>
<td>KCT</td>
</tr>
<tr>
<td>Kennedy et al. (2017)</td>
<td>M</td>
<td>&gt;yr.</td>
<td>ASS</td>
<td>Discussion</td>
<td>Coaching (peer model + feedback)</td>
<td>Resource + CS</td>
<td>KCT</td>
</tr>
<tr>
<td>Lotter et al. (2013)</td>
<td>S</td>
<td>&lt;yr.</td>
<td>SDI</td>
<td>Practice</td>
<td>Peer model + feedback, collaborative inquiry groups, &amp; TEC</td>
<td>Workshop + CS</td>
<td>KCT</td>
</tr>
<tr>
<td>Maher et al. (2014)</td>
<td>M</td>
<td>Varied</td>
<td>ASS</td>
<td>Inquiry; Discussion</td>
<td>Group problem solving &amp; TEC</td>
<td>Resource + CS</td>
<td>KCS</td>
</tr>
<tr>
<td>Patel et al. (2012)</td>
<td>M</td>
<td>&gt;yr.</td>
<td>ASS</td>
<td>Curriculum</td>
<td>No-Collaborative discussion opportunities (NOS)</td>
<td>No-Workshop only</td>
<td>KCC</td>
</tr>
<tr>
<td>Penuel (2009)</td>
<td>S</td>
<td>=1 yr.</td>
<td>ASS</td>
<td>Inquiry; Curriculum Planning Practice</td>
<td>Lesson/unit planning in teams</td>
<td>Workshop + CS</td>
<td>KCC</td>
</tr>
<tr>
<td>Singer et al. (2011)</td>
<td>S</td>
<td>&gt;yr.</td>
<td>ASS</td>
<td>Planning; Practice</td>
<td>Lesson/unit planning in teams</td>
<td>Workshop + CS</td>
<td>KCT</td>
</tr>
<tr>
<td>Walker et al. (2012)</td>
<td>M</td>
<td>&gt;yr.</td>
<td>ASS</td>
<td>Practice</td>
<td>Collaborative discussion</td>
<td>Workshop + CS</td>
<td>KCT</td>
</tr>
</tbody>
</table>
Note. CF = content focus; PCK = pedagogical content knowledge; M = math; SDI = school district initiative; KCS = knowledge of content and students; MS = math and science; ASS = active system of support; TEC = technology-enhanced collaboration; CS = collaborative support; KCT = knowledge of content and teaching; S = science; KCC = knowledge of content and curriculum; NOS = not otherwise specified.
Table 3

Sample Participant Characteristics

<table>
<thead>
<tr>
<th>Study</th>
<th>Teacher Total</th>
<th>Teaching Experience</th>
<th>Grades 6,7,8, or other (O)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0-5</td>
<td>6+</td>
</tr>
<tr>
<td><strong>CF: Math</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anderson &amp; Hoffmeister (2007)</td>
<td>19</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>Maher et al. (2014)</td>
<td>54</td>
<td>13</td>
<td>44</td>
</tr>
<tr>
<td>Patel et al. (2012)</td>
<td>57</td>
<td>11</td>
<td>23</td>
</tr>
<tr>
<td>Scheeler et al. (2010)</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>CF: Science</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hanegan et al. 2009</td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Lotter et al. (2013)</td>
<td>25</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>Penuel (2009)</td>
<td>34</td>
<td>NS(M8)</td>
<td>NS</td>
</tr>
<tr>
<td>Singer et al. (2011)</td>
<td>13</td>
<td>0</td>
<td>36</td>
</tr>
<tr>
<td><strong>CF: Math &amp; Science</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blanchard et al. (2016)</td>
<td>20</td>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td>Walker et al. (2012)</td>
<td>36</td>
<td>NS(M9)</td>
<td>NS</td>
</tr>
<tr>
<td>Kennedy et al. 2017</td>
<td>3</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td><strong>TOTAL:</strong></td>
<td>295</td>
<td>86 (34% of 250)</td>
<td>164 (66% of 250)</td>
</tr>
</tbody>
</table>

Note. NS = grades 6-8 (not specified); /-/ = not included; M = mean
Table 4

*Effect Sizes Per Study on Teacher and Student Outcomes*

<table>
<thead>
<tr>
<th>Study</th>
<th>Hedge’s g (95% CI)</th>
<th>Tau-U</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Teacher Outcomes</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scheeler et al. (2010)</td>
<td>ND</td>
<td>1</td>
</tr>
<tr>
<td><strong>Student Outcomes</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anderson &amp; Hoffmeister (2007)</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Blanchard et al. (2016)</td>
<td>.38 (−0.26, 1.01)</td>
<td>ND</td>
</tr>
<tr>
<td>Hanegan et al. 2009</td>
<td>.57 (−0.69, 1.84)</td>
<td>ND</td>
</tr>
<tr>
<td>Kennedy et al. 2017</td>
<td>1.88 (0.52, 3.24)</td>
<td>1.17 (0.89, 1.44)</td>
</tr>
<tr>
<td>Lotter et al. (2013)</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Maher et al. (2014)</td>
<td>.899 (ND)</td>
<td>ND</td>
</tr>
<tr>
<td>Patel et al. (2012)</td>
<td>.39 (0.02, 0.77)</td>
<td></td>
</tr>
<tr>
<td>Penuel (2009)</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Singer et al. (2011)</td>
<td>1.37 (0.52, 2.23)</td>
<td></td>
</tr>
<tr>
<td>Walker et al. (2012)</td>
<td>.16 (−0.3, 0.62)</td>
<td>.03 (−0.09, 0.14)</td>
</tr>
</tbody>
</table>

*Note.* CI = confidence interval; ND = not determined; *a*Study did not provide enough information to determine effect size.
Table 5:

**Quality Indicators of Group Design Studies**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Participants</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Describes Sufficiently</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Comparable across conditions</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Interventionists described</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>IV</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Clearly Described IV</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Described Fidelity of implementation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comparison condition described</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DV</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Multiple measures used</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Outcomes measured at appropriate times</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Results</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Appropriate data analysis techniques</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effect size reported</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

Desirable quality indicators
Participants
<table>
<thead>
<tr>
<th></th>
<th>X</th>
<th>X</th>
<th>X</th>
<th>X</th>
<th>X</th>
<th>X</th>
<th>X</th>
<th>X</th>
<th>X</th>
<th>X</th>
<th>X</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attrition rate &lt;30%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DV</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evidence of test-retest reliability, internal consistency reliability, and IRR</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Adequate IOA Score</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data collectors blind to study conditions</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outcomes measured beyond immediate posttest</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Criterion and construct validity provided</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Included fidelity of implementation</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comparison conditions described</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Results</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Audio or videotape excerpts included</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Results were clear and coherent</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Total indicators met** 8 11 13 15 14 13 14 12 14 16
Appendix B

Measures

Project REAL Classroom Observation

*Fill this section out ahead of time, if possible (except start and end times).*

School Code: ___________________________ Begin (time): ___________________________

Teacher Code: ___________________________ End (time): ___________________________

Observer: ___________________________ Date: ___________________________

Please check all that apply

*Fill this section out, but it is okay if you don’t get an exact count of students, or if things change partway through the observation period.*

Instructional Format:

- [ ] Whole group
- [ ] Small group
- [ ] Independent

Instructional Focus:

- [ ] Academic
- [ ] Non-academic:

# of Adults: _____

# of Students: _____

Activities/Programs Observed in Classroom

Use this space for narrative description that will help you remember events that occurred during the observation. These specifics may help you decide between 2 ratings, and may also keep you from focusing too much on a particular incident or on the end of the observation period.
<table>
<thead>
<tr>
<th>Rating</th>
<th>Behavioral Indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>(circle one)</strong></td>
<td><strong>Area of Focus – Instructional Protocols</strong></td>
</tr>
</tbody>
</table>
| 1 2 3 4 5 6 | Lesson plans match the planning board/class schedule. 

*This will generally be a 1 or a 4 – the plans match or they do not. You may want to code a 2 or 3 if the planning board or schedule is very vague or if there is a minor deviation.*

Ex. Schedule reads
- Morning routine
- Academics
- Lunch
- Academics
- Dismissal

<table>
<thead>
<tr>
<th>Rating</th>
<th>Behavioral Indicator</th>
</tr>
</thead>
</table>
| 1 2 3 4 5 6 | States clear behavioral expectations before the beginning of a new activity. 

1 - Teacher does not state behavioral expectations.

2 - Teacher states expectations, but they are unclear, contradictory, or vague.

*Ex. “I expect you to work independently on this group project.”*

*Ex. “Behave yourselves.” (This one could be okay if the students have clearly been taught what it means, but if they do not meet the teachers expectations, you may assume they don’t understand what they are supposed to do.)*

3 - Teacher states expectations before some activities, but not others.

4 - Teacher states expectations before each new activity.

5 - This is the appropriate code if the students clearly know what is expected of them for each activity because they have already been taught.
Ex. Teacher says it is time for math and the students all clear their desks and get out a white board and marker.

6 - This is appropriate if there are no transitions from activity to activity during the observation period.

<table>
<thead>
<tr>
<th>1 2 3 4 5 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uses restatement of behavioral expectations to re-establish order.</td>
</tr>
<tr>
<td>1 - Teacher does not restate behavioral expectations.</td>
</tr>
<tr>
<td>2 - Teacher restates behavioral expectations but in an incomplete and/or negative manner that does not impact student behavior.</td>
</tr>
<tr>
<td>Ex. “Stop it! Be quiet!”</td>
</tr>
<tr>
<td>3 - Teacher restates behavioral expectations, but not consistently or with inconsistent results.</td>
</tr>
<tr>
<td>4 - Teacher restates behavioral expectations consistently and students respond well. Even if only one or two instances occur, a 4 may be coded if you get the sense that this is standard procedure for the teacher.</td>
</tr>
<tr>
<td>5 - Appropriate if the teacher does not need to re-establish order and other indicators suggest that the teacher would restate expectations if needed.</td>
</tr>
<tr>
<td>6 - Use if there are no instances of behavior that need redirection, and other indicators do not provide enough information to code a 5 or do not suggest that the teacher would restate expectations if redirection was needed.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1 2 3 4 5 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Has materials and supplies ready at start of class.</td>
</tr>
<tr>
<td>1 - Teacher does not have materials and supplies ready and uses class time to gather and organize.</td>
</tr>
<tr>
<td>2 - Teacher has some materials ready but uses class time to gather others.</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Has opening/readiness activities. This will generally be a 1 or a 4 – the teacher has an opening activity or does not. You may want to code a 2 or 3 if the teacher does an opening activity but takes time to decide what it is or to organize it while students are waiting.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5 - Do not use for this item.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6 - Use if you do not observe the beginning of a class or new activity during the observation period.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maintains high time on task among students.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1 - Students spend a lot (no – we haven’t defined this) of time waiting for the teacher or their peers. There do not seem to be established choices for alternate activities if students finish early.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2 - There is some waiting time and/or only one or two students are actively engaged in activities at some points. Ex. During math class, the teacher has several students do problems on the board while the rest of the class watches.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3 - Most students are engaged throughout the observation period, but one or two finish early and do not appear to know what to do or there is a brief period where the students are waiting for the teacher.</td>
</tr>
<tr>
<td></td>
<td>1 2 3 4</td>
<td>5 6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Most students are actively engaged throughout the observation period.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Do not use for this item.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Do not use for this item.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

|   | 1 2 3 4 5 6 |
|---|---|---|
| 1 | Teacher does not review or summarize. |
| 2 | Teacher reviews and summarizes at least once, but does so in a manner that is not useful. Ex. “So now you know all about triangles.” |
| 3 | Teacher reviews and summarizes, but not consistently. |
| 4 | Teacher reviews and summarizes consistently and well. |
| 5 | Do not use for this item. |
| 6 | Do not use for this item. |

OVERALL – Instructional Protocols
Appendix C

SEALS II Scouting Report Form

Teacher ID

______________________________________________

Date ____________________________________________

Intervention Specialist

______________________________________________

What are focal issues for the functioning of this classroom? (Identify 1 or 2).

For each focal issue:

Reminders: Do not describe individual students or individual student behavior. The focus of the
observation is on general classroom focal issues that can be supported through intervention by
the classroom teacher.

1. Describe the classroom behaviors that are involved in the focal issue.

2. Identify which SEALS intervention areas are associated with the focal issue.

3. Identify and describe aspects of the classroom context that seem to support the behaviors.

4. Identify and describe aspects of the classroom context that could be drawn on to support
more productive behavior with respect to the focal issue.
Appendix D

SEALS Modules

I. Core Modules
   a. Grade 6 Theoretical Foundations
   b. Grade 7 & 8 Theoretical Foundations
   c. Social Dynamics Management: Theory
   d. Competence Enhancement Behavior Management
   e. Academic Engagement Enhancement

II. Supplementary Modules
   a. Neurodevelopment
   b. Promoting Positive Peer Cultures
   c. Social Dynamics Management: Practical
   d. Social Dynamics of Instructional Engagement
   e. Routines and Schedules
   f. Rules and Consequences
   g. Academic and Behavioral Challenges: Understanding Challenges
   h. Supporting Academic Engagement of Struggling Students
   i. Intervening with Behavior Problems in Struggling Students
   j. Working with Parents
VITA

Karen Rizzo

kleerizzo@gmail.com

EDUCATION

Ph.D.  Pennsylvania State University, University Park, PA  Aug. 2018
M.S.  Mercyhurst University, Erie, PA  Jan. 2010
B.S.  Edinboro University, Edinboro, PA  May 2002

TEACHING EXPERIENCE

College Teaching Experience

Instructor, SPLED 400  Penn State-Behrend
Inclusive Special Education Foundations: Law, Characteristics, Collaboration, Assessment, and Management  Spring 2018

Instructor, SPLED 403  Penn State-Behrend
Evidence-Based Instruction for Elementary Students in Reading, Writing, and Math  Spring 2018

Instructor, SPLED 461  Penn State-Behrend
Introduction to Autism Spectrum Disorder  Spring, 2018

Public School Special Education Teaching Experience

Special Education Teacher  North East, PA
North East Middle School  2007-2014

Klein Elementary School  Harborcreek, PA
2006-2007

Barber National Institute  Erie, PA
2004-2006

PUBLICATIONS
