Article

Teacher Education and Special Education
2018, Vol. 41(2) 140–157

DOI: 10.1177/0888406417745655
journals.sagepub.com/home/tes
SAGE

Introducing the Content Acquisition Podcast Professional Development Process: Supporting Vocabulary Instruction for Inclusive Middle School Science Teachers

Michael J. Kennedy[1], Wendy J. Rodgers[2], John Elwood Romig[1], Hannah Morris Mathews[1], and Katherine N. Peeples[1]

Abstract
Improving inclusive middle school science teachers' vocabulary instruction is a critical step toward helping those professionals support the learning needs of all students. In this Institute of Education Sciences (IES)–funded project, inclusive science teachers received professional development using elements of a multimedia-based approach called the Content Acquisition Podcast Professional Development (CAP-PD) intervention process. In this article, the authors discuss the three iterative pilot studies conducted as a part of this project and how each contributed to its development. Drawing on a theoretical framework that considers the characteristics of effective professional development, CAP-PD has three elements: (a) use of instructional videos that highlight the steps of practices and embed modeling videos to demonstrate practice, (b) preproduced instructional materials that support the implementation of evidence-based practices, and (c) feedback for teachers using data outputs from a classroom observation tool developed as a part of this project. Across studies, teachers who engaged in the CAP-PD process made important gains in terms of how much time they spent delivering evidence-based vocabulary instruction and the number of different practices implemented with fidelity. Future research should address questions of the lasting impact of this process on student learning and how to adapt the CAP-PD to help teachers do more than improve their vocabulary instruction.

Keywords
professional development, teacher preparation policy/service delivery, vocabulary instruction, science instruction, instructional practices, teacher preparation practices and outcomes, professional portfolios

The framers of the U.S. Constitution recognized the importance of scientific education when they granted Congress the power to "promote the progress of science and the useful arts." Yet, despite historical associations between science learning and civic, economic, and industrial progress (Quinn, Schweingruber, & Keller, 2012; Turner, 2008; Villanueva, 2010), science achievement in the United States consistently lags behind other developed nations (Fleischman, Hopstock, Pelczar,

[1]University of Virginia, Charlottesville, USA
[2]University of Nevada, Las Vegas, USA

Corresponding Author:
Michael J. Kennedy, Curry School of Education, University of Virginia, P.O. Box 400273, Charlottesville, VA 22904, USA.
Email: mjk3p@Virginia.edu

& Shelley, 2010). Our nation has taken notice. In the last quarter century, improving the quality of science education has regained prominence as a priority in the United States (Quinn et al., 2012). Although the National Assessment of Educational Progress (NAEP) science assessment data shows gradual increases in scientific achievement nationwide, only 41% of fourth-grade students and 37% of eighth-grade students demonstrated proficiency on the 2015 assessment (National Center for Education Statistics [NCES], 2015), suggesting that proficient achievement on a broad, comprehensive scale remains elusive.

The aforementioned statistics demonstrate that many American students are struggling with science achievement, but students with disabilities (SWD) are performing even worse. According to the results of the 2015 NAEP, SWD score significantly lower than students without disabilities, with only 18% of fourth graders with disabilities and 11% of eighth graders with disabilities meeting proficiency standards. This achievement gap is persistent and, based on trajectories from the last quarter century, unlikely to close without focused and empirically validated instructional interventions. That said, research suggests that SWD are capable of learning complex science content and achieving at levels commensurate with their peers without Individualized Education Programs (IEPs; Therrien, Taylor, Hosp, Kaldenberg, & Korsch, 2011).

This article addresses the problem of low achievement for SWD in science by exploring the development of an instructional intervention—Content Acquisition Podcast Professional Development (CAP-PD)—a multimedia professional development (PD) process designed to teach inclusive science teachers how to implement evidence-based practices for SWD. Our purpose herein is twofold: to disseminate the iterative process used in development of the CAP-PD across three pilot studies, and to outline the potential and promise of this innovative process for changing teachers' practice.

Effective Science Instruction for SWD

To improve academic outcomes in science for SWD, it is important to identify what the field of science education considers effective instruction for *all* students and the instructional practices that promote access to that same level of educational opportunity for SWD. At the heart of contemporary understandings of scientific learning are the Next Generation Science Standards (NGSS). With a three-pronged focus on core ideas, scientific practices, and cross-cutting concepts (Quinn et al., 2012), NGSS promote the practice of inquiry in the classroom. Sometimes mischaracterized as merely hands-on learning, inquiry-based instruction actually provides educational opportunities for students to engage with content knowledge by developing questions, making observations, and conducting investigations to discover scientific theory (Maroney, Finson, Beaver, & Jensen, 2003). Inquiry is less of an instructional activity and more of a cognitive process that promotes a deep understanding of the overarching concepts, models, and theories that order our world (Palincsar, Magnusson, Cutter, & Vincent, 2002).

Inquiry and SWD

Although few would argue with the merits and logic of inquiry-based instruction, there is no compelling empirical support for its use as a stand-alone practice for teaching SWD (Rizzo & Taylor, 2016). A small body of research, however, underscores that, when teachers provide appropriate curricular supports, an inquiry-based approach to science instruction benefits SWD (Mastropieri & Scruggs, 1992; Mastropieri, Scruggs, & Magnusen, 1999; Mastropieri et al., 2006; Scruggs, Mastropieri, & Boon, 1998; Therrien et al., 2011). Effective supports include using adapted texts, study guides, and graphic organizers; focusing on increasing the clarity and redundancy of critical content in curriculum and instruction; employing varied grouping structures that allow students to interact with peers; and implementing empirically validated strategies that target vocabulary knowledge teachers provide necessary structure. Across these studies, researchers found that providing structure to inquiry experiences was associated with positive academic, behavioral, and motivational outcomes for SWD (Rizzo & Taylor, 2016). More research, however, is needed, particularly around the role of students' vocabulary proficiency in successful engagement within inquiry-based tasks and related assessments (McGrath & Hughes, 2017).

The Importance of Vocabulary for Science Learning

In the content areas of the curriculum, such as science and mathematics, researchers have found that vocabulary knowledge is associated with increased achievement and that vocabulary deficits often serve as a barrier to learning (Baumann, Kame'enui, & Ash, 2003; Sáenz & Fuchs, 2002). Science curricula, in particular, are replete with technical vocabulary terms and concepts that impact students' understanding of texts and instruction (Bryant, Ugel, Thompson, & Hamff, 1999; Sáenz & Fuchs, 2002) and potentially prohibit equitable access to higher level tasks and processes (Elleman, Lindo, Morphy, & Compton, 2009; Yore, Hand, & Florence, 2004). Yet, this problem presents a fruitful opportunity for science teachers to structure learning experiences through the use of explicit vocabulary instruction (Archer & Hughes, 2011).

The subfield of science and disability has a history of instructional practices for teaching vocabulary that—when used with fidelity—result in increased achievement (Rizzo & Taylor, 2016; Scruggs, Mastropieri, & Okolo, 2008; Therrien et al., 2011). These vocabulary practices include direct instruction using student-friendly definitions (Ford-Connors & Paratore, 2015), purposeful opportunities for students to make connections between new and known vocabulary (McKeown & Beck, 2004; Nagy, 2005; Stahl, 2005), explicit instruction in the use of morphology (Graves, 2004; Harris, Schumaker, & Deshler, 2011; Jitendra, Edwards, Sacks, & Jacobson, 2004), and the use of examples and non-examples (McKeown & Beck, 2004; Stahl, 2005). These types of interventions or practices are effective because they provide explicit instruction that creates varied opportunities for students to connect known words to new words, thus building their mental maps of the terms and concepts that comprise the scientific discipline (Vaughn & Linan-Thompson, 2003). Using these instructional practices in tandem with an inquiry-based approach to learning allows science teachers—who are expected to provide inquiry-based instruction that aligns with state standards and/or the NGSS (NRC, 2012)—to structure learning experiences in ways that have the potential to mitigate the persistent achievement gap between SWD and their same-grade peers (NCES, 2015).

Present State of Science Instruction for SWD

Descriptive research on instruction in the field of science instruction for SWD is limited, but demonstrates that many science teachers' lessons focus heavily on whole group, teacher-dominated lecture, often at the expense of inquiry-based or lab-based activities (Kennedy, Rodgers, Romig, Lloyd, & Brownell, 2017; Harbort et al., 2007; King-Sears, Brawand, & Jenkins, 2014; Moin, Magiera, & Zigmond, 2009; Mutch-Jones, Puttick, & Minner, 2012). Observational studies of inclusive classrooms have confirmed that the general educator is the primary presenter of information to all students, even in co-taught classes, for whole-group lessons and small-group interactions (Harbort et al., 2007; King-Sears et al., 2014; Moin et al., 2009). These observational data support the need for general educators to understand how to implement effective practices for SWD.

One potential reason for teachers' reliance on teacher-dominated instruction is that the curricular expectations placed on general education science teachers at the secondary level are "a mile wide and an inch deep" (Kesidou & Roseman, 2002). Possibly because of the vast amount of content to be covered in preparation for state and other assessments, teachers revert to teaching methods that are perceived to be efficient (e.g., whole-group lecture) to get through the curriculum (Schmidt, Wang, & McKnight, 2005). This means the reality of science instruction for SWD is quite different from the ideal union of inquiry and structured supports researchers have identified as beneficial for SWD (Mastropieri & Scruggs, 1992; Mastropieri et al., 1999; Mastropieri et al., 1998; Scruggs et al., 1998; Therrien et al., 2011).

Need for PD for Inclusive Science Teachers

One reason for the disconnect between the type of instruction researchers have found effective for SWD and the instruction those students actually receive may be related to a gap in teacher knowledge. In the United States, the majority of SWD receive the bulk of their education in the general education setting (U.S. Department of Education, 2016), and the general education setting is often the only opportunity SWD have to access the science curriculum (Vannest et al., 2009). Consequently—though there are often opportunities for students to receive supplemental support outside of the classroom—the individual mainly responsible for science instruction is the general educator (Vannest et al., 2009). This is problematic because general education teachers often report feeling unprepared to teach SWD and unsupported in the effective use of inclusive instruction and evidence-based practices for SWD (Robinson, 2002; Wei, Darling-Hammond, & Adamson, 2010). In a recent position statement, the National Science Teachers Association (2017) recognized this problem and presented PD as one possible solution for helping science teachers "... learn about the unique needs of students with exceptionalities and how to meet those needs in the science classroom" (p. 2).

If teachers do not receive a strong grounding in teaching practices for SWD during their in-service program, it is typically left to school or district provided PD offerings to supplement their knowledge and skills. That said, PD offerings are usually expensive and time-intensive, and have questionable effectiveness (Yoon, Duncan, Lee, Scarloss, & Shapley, 2007). Researchers have offered multimedia-based PD as an alternative to traditional PD sessions (see Kennedy, Hirsch, Rodgers, Bruce, & Lloyd, 2017), but ongoing rigorous research is needed to capture its effectiveness and efficiency.

To address these concerns, our research team conducted a series of iterative studies with the CAP-PD process at the core to attend to the needs of students—through promoting the evidence-based practices highlighted in this article—and of teachers—by drawing on evidence-based principles of effective PD. In the remainder of this article, we discuss the conceptual framework driving this work and the research process through which we developed and adapted the CAP-PD.

Conceptual Framework

The design of the CAP-PD process has the cognitive apprenticeship framework (Collins, Brown, & Newman, 1989) as its foundation. Cognitive apprenticeship is a multifaceted framework for instruction that has implications for the content, methods, sequencing, and sociology of learning. At the core of cognitive apprenticeship is a model–coach–scaffold cycle to support the learning of new skills. These elements are heavily present in CAP-PD (described below) and have considerable support in the PD research literature. For example, Darling-Hammond, Hyler, and Gardner (2017) conducted a synthesis of PD studies that had positive effects on student outcomes and found that modeling was a common practice across many of these studies. Also, Kraft, Blazar, and Hogan (2017) conducted a meta-analysis of 44 studies examining the effects of coaching and found that, on average, coaching had both a moderately large positive effect on teachers' instruction and a smaller positive effect on student outcomes. Similarly, other aspects of cognitive apprenticeship are supported by research. Although cognitive apprenticeship lacks research supporting its use as a whole in teacher training contexts, the CAP-PD research base has begun to contribute experimental evidence in support of cognitive apprenticeship as a framework for designing PD opportunities for teachers.

CAP-PD Process

In response to the vocabulary needs of students with and without disabilities in middle school science and in line with our conceptual framework, we developed the three interrelated components of CAP-PD intervention:

1. use of CAPs for Teachers with Embedded Modeling Videos (CAP-TV);
2. use of instructional slides teachers can adopt and adapt for immediate use (Content Acquisition Podcasts–Teacher Slides [CAP-TS]); and
3. receipt of feedback and coaching on implementation of specific practices.

Each component is described in detail in this section. The goal is to create a PD intervention process that supports teachers' use of evidence-based practices for teaching critical vocabulary terms and concepts to students in content area classrooms in a way that results in meaningful learning outcomes for teachers and students. In addition, the PD process should be multimedia, such that it is sustainable, scalable, and possible to replicate. The studies in this article specifically address middle school science classrooms.

CAP-TV

CAP-TVs are short (10-15 minutes), multimedia vignettes that (a) provide direct instruction about specific vocabulary practices (CAPs for Teachers [CAP-T]) and (b) model implementation with fidelity (CAP-TV). These videos are designed in accordance with Mayer's (2009) Cognitive Theory of Multimedia Learning and consist of clear images, limited text, tight narration, and a short modeling video. A sample CAP-TV is available at https://vimeo.com/143387419. Kennedy and his colleagues have completed four direct empirical tests of the impact of CAP-TV on teacher learning across content areas (Alves et al., 2017, Ely, Kennedy, Pullen, Williams, & Hirsch, 2014; Kennedy, Rodgers, et al., 2017; Kennedy, Hirsch, et al., 2017). In each study, teachers who learned from the CAP-TV approach implemented more elements of specific evidence-based practices with fidelity than comparison teachers who learned using traditional lectures or other PD approaches. CAP-TV addresses the need to focus PD on specific instructional practices rather than content more generally (Garet, Porter, Desimone, Birman, & Yoon, 2001). The CAP-TV library provides the modeling phase of the cognitive apprenticeship conceptual model.

CAP-TS

To scaffold teachers' implementation of the target instructional practices, the package includes PowerPoint slides teachers can download and use during instruction. Our research team hired middle school science teachers to write and validate the content contained within the slides. Each CAP-TS contains slides for at least one but often two or three related vocabulary terms or concepts at a time, and includes the following instructional sequence: (a) reviewing key background knowledge, (b) providing a student-friendly definition, (c) teaching relevant examples or non-examples (when appropriate), (d) highlighting and teaching morphological features of the term (when appropriate), (e) comparing and contrasting the term with semantically related terms, and (f) repeating the definition. The slides conform to Mayer's (2009) applied theoretical model in terms of how they leverage vivid images, occasional on-screen text, and a script for supporting teachers use of clear, consistent language. In line with the research regarding effective science instruction for SWD (Mastropieri & Scruggs, 1992; Mastropieri et al., 1999; Mastropieri et al., 1998; Scruggs et al., 1998; Therrien et al., 2011), the CAP-TS narratives emphasize critical content by providing a structure that is clear and redundant where appropriate, effectively pacing instruction through the use of student cues (e.g., "Let's get ready to look at some examples and non-examples of symbiosis."), and embedding opportunities to respond and receive feedback.

The instructional practices taught and modeled within the CAP-TV vignettes are embedded into the CAP-TS providing tacit reinforcement and prompting for teachers to use the practices. In fact, the model teacher uses CAP-TSs in the embedded video. To ensure the content of the PD aligns well with the content teachers are actually teaching, the CAP-TS materials are customizable and teachers are encouraged to combine them with their existing instructional materials. In particular, the CAP-TS slides were created using vocabulary lists provided by teachers and a review of the state science standards, thereby supporting the material teachers are required to teach. The CAP-TS illustrate the modeling phase of the cognitive apprenticeship conceptual framework, and provide scaffolding as well. Sample CAP-TS can be found at www.VocabSupport.com.

Feedback From Observations and Coaching

The final component of CAP-PD is to observe teachers' instruction using an observation instrument called the Classroom Teaching (CT) Scan, which is a low-inference instrument that records instructional moves in real time (Kennedy, Rodgers, et al., 2017. The CT Scan allows for collection of data concerning the duration of discrete instructional practices used as well as counts of questions and feedback statements teachers provide to students. The CT Scan also records the quality of implementation of the practice, or the extent to which teachers' enactment of the practices corresponds to the instruction provided in the CAP-TVs. The CT Scan produces visual graphic outputs representing the use of class time and the practices used, which are provided to teachers along with instruction about how to interpret the graphs (the CT Scan can be found at www.classroomteachingscan.com/ctscan/ and a sample output can be seen at http://www.classroomteachingscan.com/ctscan/timeline.htm?menus.txt&213).

To support the visual feedback, teachers receive an email highlighting specific aspects of their instruction (see http://www.classroomteachingscan.com/ctscan/reportmaker.htm?213, for a customizable online template). The email can include specific praise for things the teacher did well (e.g., "You provided your students an anticipatory set to let them know what to expect from this class, and you did it beautifully, with a cue to get their attention and a reference to the relevance and importance of the topic at hand."). The email can also provide specific suggestions for improvement (e.g., "I noticed that during the last 30 minutes or so of the class, you only asked 4 questions. Asking more questions or giving students other ways to respond to the content, particularly at the close of a lesson, helps to cement their learning and makes it more likely that they will remember the information."). Within the template, researchers send teachers "One Big Thing," which is a key area for them to work on prior to the next lesson or observation. The observations and related feedback add to the active learning component, tying the PD directly to teachers' instruction and making them part of the learning process. Using the CT Scan and resulting descriptive data for delivering coaching and feedback rounds out how the CAP-PD process operationalizes the cognitive apprenticeship conceptual framework.

Description of Research Process and Key Findings

The CAP-PD research project was funded by the National Center for Special Education Research's (NCSER) Early Career Development and Mentoring research training program. The studies conducted over the past several years within this Goal 2, Development and Innovation project, have been informed by Hill, Beisiegel, and Jacob's (2013) recommendations for conducting more rigorous research on teacher PD. In particular, the CAP-PD studies incorporate "a series of small-scale but rigorous trials to determine the effects of various program delivery features on mostly proximal outcomes" (p. 485) in the early stages of PD development, with the intent of creating a PD intervention ready for further efficacy trials in a subsequent project.

Thus far, the first two stages of Hill and colleagues' (2013) recommended progression have been used in the CAP-PD research program to help refine the CAP-PD in preparation for future studies (i.e., "one-site pilot" studies and "randomized controlled trial holding content the same but varying features of program delivery"; p. 479). The questions driving this research are as follows:

Research Question 1: What are teachers' initial perceptions of the CAP-PD process and its perceived usefulness?
Research Question 2: For researchers, in what ways is the CAP-PD process feasible to implement?
Research Question 3: To what extent does engagement in the CAP-PD process result in inclusive middle school science teachers making measurable gains in the quality and quantity of their vocabulary instruction?
Research Question 4: Do SWD make learning gains when taught by teachers using high quality vocabulary instruction?

Teacher and Student Population Studied

A total of 34 teachers drawn from 18 rural schools participated in three studies spanning the 2014-2017 school years (three teachers in Pilot Study 1 [2014-2015], three in Pilot Study 2 [2015-2016], and 28 in Pilot Study 3 [2016-2017]). The teachers in our studies were middle school science teachers who worked in inclusive classrooms, defined as classes comprising students with and without disabilities. All teachers worked in rural districts that served students from predominantly lower socioeconomic backgrounds. Of the 36 teachers, all were certified to teach science in their state, and the average number of years teaching was 11.2 (*SD* = 3.6). Six of the teachers were male, one was Hispanic, and one was Asian. The focus to date has been on teacher performance and outcomes; however, 1,779 total students participated in the third study, of which 251 (14.1%) were identified as having an IEP. Observations of teacher practice were conducted live and via video.

Pilot Study 1

In the first pilot study, three middle school science teachers from different schools received an early version of three CAP-Ts (contained narration and images only, no embedded modeling video) to learn three key practices for teaching vocabulary: (a) use of student-friendly definitions, (b) examples and non-examples, and (c) the highlighting and defining of morphological parts of words. Each teacher watched each video once. The teachers had been teaching for 3, 5, and 7 years. The teachers also received early versions of CAP-TS to support implementation of practices. The early-version CAP-Ts and CAP-TS, however, were not perfectly aligned (i.e., the practices taught within the CAP-Ts were neither broken into specific steps nor specifically represented in the slides). The assumption was that exposure to the CAP-Ts and slides would be enough to spur changes in implementation of the three vocabulary practices. At the time of this study, the CT Scan did not exist in its current form. A paper-based predecessor was used to record data but teachers did not receive feedback on their performance using the CT Scan as part of the intervention. The main purpose of this study, in accordance with Stage 1 of Hill et al.'s (2013) recommended progression, was to determine whether the PD process was feasible to implement and to collect social validity data from participating teachers.

Method. The teachers videotaped their teaching 5 days prior to accessing the CAP-Ts (baseline), and then 5 days after (post intervention). A one group, pre–post design was utilized. Teachers uploaded videos to a secure collaborative site provided by the researchers' university. Two members of the research team coded approximately 25% of the videos. The observers were the principle investigator and two doctoral students, each with an average of 6 years as classroom teachers and are the authors of the

CT Scan. Inter-scorer reliability was 70% for number of vocabulary practices being used by teachers. Based on these results, members of the research team noted patterns across lessons and compared impressions of the instruction. Data were not analyzed in relation to PD effectiveness, but on teachers' use of the CAP-TS and their perceptions of the process.

Results. Prior to the CAP-PD, teachers provided textbook definitions or had students copy words and definitions from the glossary in their text during each lesson (teachers recorded lessons on days when they taught vocabulary). After viewing the PD videos, teachers used more vocabulary practices. After the intervention, teachers provided a student-friendly definition in each lesson, and provided examples and non-examples for each term, as outlined in the CAP-TS. Researchers measured use of these practices using the predecessor for the CT Scan. We did note that teachers did not seem to take ownership of the CAP-TS, instead using them as packaged curriculum. Teachers displayed the slides to students and read what was on them, but they did not alter the slides to align them with the rest of their instruction. There was also some indication that teachers did not even review the CAP-TS before class—they would often pause when a slide came up before making a statement such as follows: "We're not going to talk about that word today—let's move to the next one."

After this pilot study was completed, we communicated with participating teachers by email to determine their impressions of the CAP-PD. All three teachers in their written comments to researchers revealed a desire for a more intensive PD option than simply watching three short videos and accessing slides, and for timely feedback on their teaching. A representative quote was, "I appreciate having access to the videos and slides, but I didn't really know what I was supposed to do with them. I did the best I could." A second teacher wrote, "I filmed myself like you asked, but I didn't know if I was doing what you wanted. Next time it would be good to get some feedback."

Iterative development of PD materials. This study served as a feasibility test, in accordance with the first stage of the recommended sequence for PD research proposed by Hill and colleagues (2013). Recommendations of the teachers for more robust training spurred the development of the CAP-TV library, which was upgraded from the original CAP-Ts to include specific steps of implementation for each practice and the inclusion of a modeling video. Experts in the field of vocabulary instruction reviewed each draft CAP-TV and provided input on adherence to best practice. In addition, as a result of our observations that teachers did not take ownership of the CAP-TS, we altered them to be more aligned with the CAP-TV.

In addition, in response to the teachers' request for feedback, a computer-based version of the CT Scan was developed that generated immediate data to use as part of a coaching and feedback cycle. The revised CT Scan also allowed for more specific information to be collected on the implementation of the vocabulary practices. Previously, we had developed lists of steps we expected teachers to follow when implementing the practices. The paper version afforded an assessment of implementation fidelity using a binary system (i.e., yes or no), but the specific implementation markers for each practice were not listed on the CT Scan data tracking form—observers would have to use the codebook or have the markers memorized (which proved very difficult when using a real time, low-inference tool such as the CT Scan).

The computer-based CT Scan improved the evaluation by providing the observer with a list of implementation markers for each practice. The implementation markers are drawn from the professional literature for each practice, when available. For example, the implementation markers for having a discussion are drawn from Ford-Connors and Paratore's (2015) synthesis of the literature. The markers are as follows: call on students all over the room, encourage students to talk and use evidence, provide an appropriate and authentic context, ask open-ended questions, and incorporate student responses into discussions. As an observer watches a teacher use this practice, they use the CT Scan to click on each implementation marker as a guide, and those left unclicked become the targeted areas for improvement in the coaching phase. This allowed us to provide more detailed feedback following the observations and connect the feedback directly to the CAP-TV and CAP-TS components. These implementation markers also greatly improved the observers' interscorer reliability, and provided content validity as well (see below).

Pilot Study 2

Due to the extensive changes made to CAP-PD intervention after conducting Pilot Study 1, a Second Stage 1 feasibility study was conducted (Hill et al., 2013). Pilot Study 2, which now included the CAP-TV, integrated with the CAP-TS, and coaching and feedback. A single-case, multiple-baseline design study was utilized to experimentally measure the impact of the CAP-PD intervention on inclusive science teachers' implementation of evidence-based vocabulary practices. Three teachers with 2 to 6 years of teaching participated in this study. These teachers had a total of 26 students with IEPs spread across their classes. Observations took place for approximately 6 weeks, and teachers were observed between 18 and 24 times between the baseline, post intervention, and maintenance observations.

Method. Following baseline observations, teachers received access to five CAP-TV (student-friendly definitions, examples and non-examples, morphological parts, semantic relationships among terms, and having discussions), which presented specific steps for each practice that were derived from the empirical literature on vocabulary instruction and explicit instruction. Each teacher watched each video at least once, and further viewings were tracked by researchers using www.edpuzzle.com. Next, teachers received access to CAP-TS for each vocabulary term in their curriculum that were closely aligned with the CAP-TV. Finally, after each observation in the intervention phase, teachers received a daily email containing data output from the CT Scan and a personalized email providing feedback and coaching notes as described above. When improvement was needed, the email would refer the teacher back to a specific point in a CAP-TV to provide a model of or information about proper implementation.

Observations were conducted live, and a second observer was present for 25% of the observations (85% agreement on amount of time spent teaching vocabulary and specific vocabulary practices). Following the observations, the two observers compared results and came to consensus on what had occurred. Data on the target outcomes (i.e., amount of time spent in explicit vocabulary instruction, number of vocabulary practices, and percentage of implementation markers used correctly within each practice taught in the CAP-TV) were graphed and visually analyzed for significance.

Results. At baseline, teachers spent on average 10% of class time (about 5 minutes per day) using evidence-based vocabulary practices. The most vocabulary practices used with fidelity by any teacher were two for any given lesson. Following intervention, on average, teachers spent 46% of class time on vocabulary instruction and used 4.6 practices with fidelity per lesson. In the course of the entire study, only two lessons from Teacher C contained instruction that could be considered inquiry-based. At baseline, teachers were coded as "observing and assessing" students, on average, 42% of the time, which indicates the amount of time students were working on independent or small-group assignments while teachers monitored them. These assignments were not inquiry-based and instead were lower level activities such as students working in groups to create study guides for vocabulary terms. This percentage was reduced to less than 20% post intervention, with the lower level activities being replaced with explicit instruction provided using the CAP-TS and other instruction promoted by the CAP-TV. The CAP-PD intervention process was thus successful not only in improving teachers' vocabulary instruction but also in reducing the number of lower level activities in which students engaged. Full methods and results for this study are available in Kennedy, Rodgers, et al., 2017.

Iterative development of PD materials. During observations, researchers noted that teachers took ownership of the CAP-TS by customizing them to enhance their instruction. All three teachers used the slides as a foundation and made a number of changes to seamlessly incorporate materials into their practice. Some instructional changes were substantive, including the addition of videos, diagrams, and demonstrations. Other changes were less so and focused on visual enhancements, such as changes in the colors, fonts, or size of images to match existing materials. Both substantive and design-related changes allowed teachers to use the materials with a greater level of comfort than was observed in Pilot Study 1.

Although the CAP-PD improved teacher practice, the researchers' workload was heavy. Teachers were observed daily and provided feedback after each observation. This high frequency of observations is not sustainable for long periods of time and is not scalable to large numbers of teachers. Therefore, the next study focused on teasing apart the additive effects of each CAP-PD component and on determining whether teachers' practices could be improved while maintaining reasonable demands on the PD providers.

Pilot Study 3

In our third study, we moved to Stage 2 of Hill et al.'s (2013) PD research stages by testing the impact of specific PD features on teachers' practice. Twenty-eight inclusive middle school teachers from 11 rural schools agreed to be randomly assigned to one of two experimental conditions. Although underpowered, a randomized control trial design was utilized. All teachers were observed 3 times to establish a baseline measure of their vocabulary instruction, and then 3 more times following exposure to PD materials at an interval of approximately once per month. Teachers in the treatment condition received the full CAP-PD intervention process, including access to the five CAP-TV, www.vocabsupport.com, a website containing approximately 100 CAP-TS, and coaching emails using data outputs from the CT Scan after each of the three post intervention observations. Teachers in the comparison condition received access to the CAP-TS only, without the CAP-TV or feedback. Researchers therefore tested the impact of the CAP-TV and coaching emails using the CT Scan while holding access to the CAP-TS constant. Each teacher in the CAP-PD group watched each CAP-TV twice, per data collected by www.edpuzzle.com.

Method. Members of the research team who conducted observations were kept blind as to which group the teachers belonged. Teachers were given the CAP-PD materials according to their assigned condition after their third baseline observation. Most observations were conducted live, but due to scheduling and travel constraints, several had to be recorded and coded from video. For 20% of the in-person observations, a second observer was either present in the classroom or watched the video. Inter-observer reliability was calculated by comparing percentages on the target outcomes. Inter-observer agreement for percent of time spent teaching vocabulary was 95%, use of specific vocabulary practices was 92%, and agreement on implementation markers within each vocabulary practice was 89%. Disagreements were resolved via discussion.

Goal 2 studies support the development of interventions intended to produce positive outcomes for students in authentic settings, such as science classrooms. Consequently, student-level data were also collected in Pilot Study 3—1,779 students participated, including 251 with IEPs. Over 70% of all students were from low socioeconomic backgrounds as measured by their receipt of free or reduced lunch at school. Students completed a standardized pre- and posttest of science knowledge (e.g., Misconceptions-Oriented Standards-Based Assessment Resources for Teachers [MOSART], see https://www.cfa.harvard.edu/smgphp/mosart/testinventory_2.

html), and three curriculum-based vocabulary probes (20 multiple-choice terms each, randomly drawn from the teachers' curriculum). The curriculum-based measurement (CBM) probes were spaced to coincide with monthly observations, and confirmed in structure and procedures for administration to best practice using CBMs described by Espin, Shin, and Busch (2005) and others.

Quality vocabulary index (QVI). The CT Scan does not automatically generate a quality score following an observation. In Pilot Study 2, simple descriptive data were used to examine differences as the teachers progressed through the CAP-PD process. To analyze, however, descriptive data that the CT Scan does collect in a randomized control trial design, it was necessary for researchers to have an interpretable score. Therefore, the QVI was created.

The QVI is a new, but transparent statistic that allows researchers to use descriptive data from teachers' observations and make intergroup comparisons, and evaluate growth over time. The QVI is presented mathematically below:

$$\text{QVI} = \Sigma\left(\left(x+1\right)\frac{y}{z}\right)_{ij},$$

where x is the percentage of implementation markers (fidelity) used for a given practice, y is the duration (seconds) that the teacher demonstrated the given practice, and z is the duration (seconds) of the entire lesson. Each $(x+1)y/z$ is the percentage of implementation markers used for an individual practice weighted by the proportion of class time spent in that practice. This term is calculated for each distinct iteration of a practice. If a teacher uses three separate practices (student-friendly definition, morphological analysis, and examples of terms) in a lesson, she would have three weighted fidelity terms. Likewise, if the teacher used one practice (examples of terms) three separate times in a lesson, she would have three weighted fidelity terms. All weighted fidelity terms for each practice in an observation are summed to create the total QVI score for an observation.

One assumption we made about instruction is that the QVI should give teachers credit for using a practice even if they demonstrated 0% fidelity. Our rationale was that using an evidence-based practice poorly should be better than not using evidence-based practices at all. Because of this, we added a "1" to every x. If we had not done this step, the score would not have included any practices where the teachers demonstrated 0% fidelity because 0 multiplied by anything else would equal 0. Table 1 includes two examples of QVI scores based off actual teacher data. The first represents a high example, and the second is a low example. The second example in Table 1 demonstrates that even if a teacher uses two evidence-based practices with high quality, the resulting QVI will be low if not used for much time (less than 3/90 minutes in this example).

Teacher results. Researchers conducted 76 baseline and 80 post intervention observations for a total of 8,767 minutes (average = 55.8 minutes/lesson) between August and January of an academic year. At baseline, teachers in the treatment condition (full CAP-PD) spent an average of 8.2 minutes (SD = 13.4) explicitly teaching vocabulary per lesson, and had a mean QVI of .2460, SD = 0.321. Colleagues in the CAP-TS-only condition explicitly taught vocabulary for an average of 9.5 minutes (SD = 10.3) per lesson, and had a QVI of .2724, SD = 0.301, $F(1, 75) = .138$, $p = .711$. For the three observations at post intervention, teachers in the CAP-PD group spent an average of 29.5 minutes per lesson explicitly teaching vocabulary (SD = 13.5), and saw their mean QVI rise to .9035, SD = 0.377. Teachers in the CAP-TS group spent an average of 15.9 minutes per lesson explicitly teaching vocabulary (SD = 11.6), and had their mean QVI rise to .5117, SD = 0.465. The difference between the groups' performance at postintervention was statistically significant, and a large effect size was detected—$F(1, 78) = 17.2$, $p < .001$, $d = .93$.

SWD results. SWD with teachers in the CAP-PD condition (N = 132, M = 17.8, SD = 6.9) did

Table 1. Sample QVI Data.

Teacher	Duration (minutes)	Vocabulary practice	Duration of practice (seconds)	Percentage of fidelity implementation	QVI
082	90	Student-friendly definition	40	0.33	1.49
		Formal definition	66	1.00	
		Example	558	1.00	
		Ask students to state definition	18	0.50	
		Example	481	1.00	
		Student-friendly definition	111	1.00	
		Example	200	1.00	
		Ask students to state definition	28	1.00	
		Morphological analysis	46	0.75	
		Student-friendly definition	98	1.00	
		Example	260	1.00	
		Student-friendly definition	36	1.00	
		Morphological analysis	55	0.25	
		Example	122	0.75	
		Semantic feature analysis	122	0.66	
		Example	229	1.00	
		Student-friendly definition	189	1.00	
		Example	663	0.50	
		Student-friendly definition	73	1.00	
		Morphological analysis	24	0.50	
		Student-friendly definition	39	0.33	
		Analogy	102	1.00	
		Example	362	1.00	
		Student-friendly definition	120	1.00	
061	90	Ask student to state definition	43	0.50	0.07
		Example	106	1.00	

Note. QVI = quality vocabulary index.

not score significantly differently than peers with teachers in the CAP-TS-only condition ($N = 119$, $M = 18.5$, $SD = 6.7$) on the MOSART pretest, $F(1, 249) = .517$, $p = .473$. *SWD*, however, in the CAP-PD group ($M = 23.3$, $SD = 7.4$) did significantly outperform peers in the CAP-TS-only group ($M = 20.1$, $SD = 8.0$) on the MOSART posttest, $F(1, 238) = 9.8$, $p = .002$, $d = .54$. The MOSART has three parts: physical science, life science, and astronomy, and has 60 total points.

SWD with teachers in the CAP-PD condition (CBM 1, $M = 10.3$, $SD = 3.5$; CBM 2, $M = 11.9$, $SD = 3.4$; CBM 3, $M = 13.0$, $SD = 3.5$) significantly outscored peers with disabilities from the CAP-TS condition (CBM 1, $M = 9.1$, $SD = 3.5$; CBM 2, $M = 9.3$, $SD = 3.6$; CBM 3, $M = 10.4$, $SD = 3.5$) on all three CBM probes: CBM 1, $F(1, 240) = 6.5$, $p = .011$, $d = .33$; CBM 2, $F(1, 239) = 31.5$, $p < .001$, $d = .73$; CBM 3, $F(1, 228) = 28.2$, $p < .001$, $d = .70$. Each CBM probe had 20 vocabulary terms randomly drawn from the teachers' curriculum and was administered with procedures adapted from Espin et al. (2005).

Iterative development of PD materials. These results demonstrate that having access to five CAP-TV and three feedback emails using the CT Scan provides teachers with enough information regarding their instructional performance to change their practices, compared with the CAP-TS only. In addition, the SWD from teachers' who participated in the CAP-PD made important and statistically significant gains in their performance on a standardized science measure, and three CBM probes. In future studies of the CAP-PD intervention process, we plan to maintain the reduced number of observations rather than attempting to conduct them daily.

In addition, data were gathered from participants using an online survey that informed updates to the PD materials. First, teachers reported benefiting from the CAP-TVs, CAP-TS, and CT Scan feedback emails (Prompt: The quality of my teaching benefited from participating in this project: strongly agree, agree, neutral, disagree, strongly disagree). The mean score for teachers in the CAP-PD condition was 4.2 (SD = 0.3).

The survey concluded with an open field for suggestions. Five teachers proposed collapsing the CAP-TS into groupings by topic unit (e.g., individual slide shows on velocity, speed, motion, and acceleration were combined into one, as these terms are often taught together). Four teachers noted liking the CT Scan data and feedback, but were overwhelmed at first stemming from how different it is from any feedback previously received. The research team responded by making a tutorial for future teachers on how to interpret the CT Scan data. Teachers also requested that the CAP-TS include more than vocabulary instruction. They asked for sample demonstrations and hands-on activities to be included within the slides to help them connect the supports provided by explicit vocabulary instruction to higher level application activities. These changes are also present in the current versions of our curriculum materials (see www.vocabsupport.com).

Implications for Future Research

The goal of any Goal 2 Institute of Education Sciences (IES)–funded project is to iteratively develop an intervention with the intent of generating promising empirical results and to justify further experimental trials at increasingly larger scales. In the case of this project, combining an IES Goal 2 project with an Early Career Development and Mentoring project provided resources that have been pivotal in shaping and reshaping an intervention aimed at improving teacher practice. The theory and practice underlying this research builds on experimental trials conducted by Mastropieri, Scruggs, and colleagues (i.e., Mastropieri et al., 1999; Mastropieri et al., 1998; Mastropieri et al., 2006) and by researchers in the field of evidence-based practice for vocabulary instruction to build a PD process that addresses the need to build general education teachers' knowledge about and implementation of practices to support the learning of SWD in the content areas. Studies of CAP-PD, however, add a needed focus on teacher practice.

The work by Mastropieri, Scruggs, and colleagues highlighted the value of a structured inquiry approach to science education for SWD, but their research was primarily focused on manipulating the curriculum as the independent variable (Mastropieri et al., 1999; Mastropieri et al., 1998; Mastropieri et al., 2006). An example is their work with the keyword mnemonic strategy for learning the definition of individual vocabulary terms. Although the keyword mnemonic strategy can and should be included as part of teachers' instructional routines, it is not intended to be a transformative instructional practice that infiltrates all areas of instruction in science. Therefore, although teachers in Mastropieri and colleagues' studies were responsible for implementing curriculum, the focus of their body of work was not on changing teacher practice. The researchers' focus on curriculum alone overlooked the influence of teachers' knowledge and skill related to evidence-based practice. It cannot be inferred from this work that—beyond the units of study included in these experiments—there was a change in teachers' practice. Because CAP-PD studies use elements of teacher practice as the dependent variables, the impact of a PD intervention on teachers' instruction is revealed.

Reflection on Results

PD offerings in science and beyond are typically delivered using a "traditional" in-person model. This includes a trainer/expert of a new program or practice delivering content to teachers during one or several sessions of varying lengths. Once the session ends, the teachers are largely left to their own devices to implement the new practice(s) or program, and receive little, if any, feedback. Researchers across the past 20+ years have documented the inefficiency and ineffectiveness, overall, of this approach (Yoon et al., 2007). Multimedia offers an intriguing alternative to the traditional trainer/trainee model of PD; however, critical questions of implementation fidelity, dosage of PD content, and the appropriateness of match between content being learned and the multimedia delivery platform must all be explored and tested experimentally (Kennedy, Rodgers, et al., 2017).

Results of the three pilot studies described in this article demonstrate that the amount of time and the number of evidence-based vocabulary practices that inclusive middle school science teachers incorporate into instruction can be increased without intensive face-to-face PD offerings. By using simple multimedia (i.e., CAP-TV and CAP-TS) and sending a modest number of feedback emails using detailed descriptive data, teachers across the two studies made measurable gains in teaching performance. In addition, preliminary evidence now exists for the impact on learning of SWD. Other than the one-time cost of authoring and producing the CAP-TV and CAP-TS, no further investments are required, and teachers can use these resources in perpetuity. Using the CT Scan to provide coaching does require an ongoing time commitment from a researcher, administrator, coach, or another professional to observe instruction either live or via video to generate feedback. Future flexible use of the CT Scan, however, as a part of regularly scheduled cycles of pre-service and in-service teacher observation and feedback could become a reality.

Notes on iterative process. The iterative development of this project across the past several years holds important lessons for other researchers and teams seeking to undertake a similar project. A key lesson learned is the need to intensify PD offerings as much as possible. This includes ensuring curriculum materials are a good match for teachers' needs, and providing intense and ongoing coaching to support implementation. A second lesson is not being satisfied with positive and statistically significant results. The second and third studies introduced in this article resulted in desired outcomes by way of positive teacher gains. The researchers, however, did not declare victory and pronounce the PD intervention as being finalized. Instead, the team scoured teachers' implementation, gathered informal feedback from participants, and made needed updates to further improve the process. Next, the decision to create an observation tool to serve as a key dependent measure for the project paid dividends and caused headaches. In other words, the iterative process of developing the CT Scan has been a project within the project. Researchers and teams should be very thoughtful and thorough when it comes to selecting the dependent measures for their projects, and in the instance when creation of something new is needed, be sure the team possesses the expertise and time needed for such an undertaking.

Directions for Future Research

Several lingering questions remain to be answered in future validation projects of the CAP-PD intervention process. First, further study of the individual components of CAP-PD is indicated. For example, the CAP-TV feature could be enhanced by embedding questions, including text overlays on the modeling videos, and other video animations. The degree to which these elements enhance learners' attention or distract from the desired learning experience deserves further study.

In terms of teachers, questions remain about how to support teacher interpretation of the CT Scan feedback and the most efficient method(s) for maximizing the impact of the feedback given. Although we have collected social validity data regarding perceived benefits of the CT Scan feedback, currently no data exist addressing the degree to which teachers can read and interpret the CT Scan outputs. It is possible the outputs could be overwhelming and that limits need to be placed on the amount and kind of information provided to teachers. Also, what are the best approaches for communicating to, and motivating, teachers based on the data outputs? Currently, narrative feedback emails are written describing the visual output and directing teachers' attention to aspects of the visual graph, but there may be more effective and motivating ways to present this information. Incorporating goal setting by having teachers review their baseline data and set goals for the number of (a) practices they want to use in a lesson, (b) questions or feedback statements they want to provide in a lesson, and/or the (c) amount of time they want to spend in an instructional domain could be an appropriate strategy for motivating teachers to improve practice and one to explore in the future.

Finally, once the development and refinement of CAP-PD in middle school science classrooms is completed, research could be extended into other grade levels and content areas. Many other secondary content areas have high vocabulary demands (e.g., history, mathematics). It is possible that teachers from disciplines other than science would respond differently to the CAP-PD package.

Conclusion

In general, results of the three pilot studies suggest that teachers understand and can use the CAP-PD, and that it is feasible to implement during typical science classroom instruction. The studies also reveal the CAP-PD to be a promising intervention to support inclusive middle school science teachers' use of research-based vocabulary practices. The model–coach–scaffold cycle used across studies for helping teachers change their practice also provides an operationalization of the cognitive apprenticeship conceptual framework. In continuing work in this domain, we aim to further explore how this conceptual framework is appropriate for iteratively designing a PD package for use with multimedia-based products. Researchers doing work in this domain can take lessons from our experience, in terms of gathering qualitative and quantitative data from small numbers of teachers and making incremental improvements to instructional materials. In addition, across these studies, researchers were continuously reminded of the need to make PD materials more intense, and specifically linked to teachers' work. This resulted in our pairing of CAP-TV and CAP-TS curriculum materials, and the design and shaping of the CT Scan.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/ or publication of this article: The research in this article was supported by an Early Career Research and Mentoring grant awarded by the National Center for Special Education Research within the Institute for Education Sciences (Cooperative Agreement # R324B130023 to Michael Kennedy at the University of Virginia. No specific endorsement of the products reported herein on the part of the U.S. Department of Education should be inferred.

References

Archer, A. L., & Hughes, C. A. (2011). *Explicit instruction*. New York, NY: Guilford Press.

Alves, K. A., Kennedy, M. J., Kellems, R. O., Wexler, J, Rodgers, W. J., Romig, J. E., & Peeples, K. N. (2017) Improving preservice teacher vocabulary instruction: A randomized controlled trial. *Teacher Education & Special Education*. Advance Online Publication. doi: 10.1177/0888406417727044

Baumann, J. F., Kame'enui, E. J., & Ash, G. E. (2003). Research on vocabulary instruction: Voltaire redux. In *Handbook of research on teaching the English language arts* (2nd ed., pp. 752-785).

Bryant, D. P., Ugel, N., Thompson, S., & Hamff, A. (1999). Instructional strategies for content-area reading instruction. *Intervention in School and Clinic*, *34*, 293-302. doi:10.1177/105345129903400506

Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glasser* (pp. 453-494). Hillsdale, NJ: Lawrence Erlbaum.

Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). *Effective teacher professional development*. Palo Alto, CA: Learning Policy Institute.

Elleman, A. M., Lindo, E. J., Morphy, P., & Compton, D. L. (2009). The impact of vocabulary instruction on passage-level comprehension of school-age children: A meta-analysis. *Journal of Research on Educational Effectiveness*, *2*, 1-44. doi:10.1080/19345740802539200

Ely, E., Kennedy, M. J., Pullen, P. C., Williams, M. C., & Hirsch, S. E. (2014). Improving instruction of future teachers: A multimedia approach that supports implementation of evidence-based vocabulary practices. *Teaching and Teacher Education*, *44*, 35-43. doi: 10.1016.j.tate.2014.07.012

Espin, C. A., Shin, J., & Busch, T. W. (2005). Curriculum-based measurement in the content areas: Vocabulary matching as an indicator of progress in social studies learning. *Journal of Learning Disabilities*, *38*, 353-363.

Fleischman, H. L., Hopstock, P. J., Pelczar, M. P., & Shelley, B. E. (2010). *Highlights from PISA 2009: Performance of U.S. 15-year old students in reading, mathematics, and science literacy in an international context* (NCES 2011-004). Retrieved from http://nces.ed.gov/pubs2011/2011004.pdf

Ford-Connors, E., & Paratore, J. R. (2015). Vocabulary instruction in fifth grade and beyond: Sources of word learning and productive contexts for development. *Review of Educational Research*, *85*, 50-91.

Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal*, *38*, 915-945.

Graves, M. F. (2004). Teaching prefixes: As good as it gets? In J. F. Baumann, & E. J. Kame'enui (Eds.), *Vocabulary instruction: Research to practice* (pp. 81-99). New York, NY: Guilford Press.

Harbort, G., Gunter, P. L., Hull, K., Brown, Q., Venn, M. L., Wiley, L. P., & Wiley, E. W. (2007). Behaviors of teachers in co-taught classes in a secondary school. *Teacher Education and Special Education*, *30*, 13-23.

Harris, M. L., Schumaker, J. B., & Deshler, D. D. (2011). The effects of strategic morphological analysis instruction on the vocabulary performance of secondary students with and without disabilities. *Learning Disability Quarterly*, *34*, 17-33.

Hill, H. C., Beisiegel, M., & Jacob, R. (2013). Professional development research: Consensus, crossroads, and challenges. *Educational Researcher*, *42*, 476-487. doi:10.3102/0013189X13512674

Jitendra, A. K., Edwards, L. L., Sacks, G., & Jacobson, L. A. (2004). What research says about vocabulary instruction for students with learning disabilities. *Exceptional Children*, *70*, 299-322.

Kennedy, M. J., Rodgers, W. J., Romig, J. E., Lloyd, J. W., & Brownell, M. T. (2017). Effects of a multimedia professional development package on inclusive science teachers' vocabulary instruction. *Journal of Teacher Education*, *68*, 213-230. doi: 10.1177/0022487116687554

Kennedy, M. J., Hirsch, S. E., Rodgers, W. J., Bruce, A., & Lloyd, J. W. (2017). Supporting high school teachers' implementation of evidence-based classroom management practices. *Teaching and Teacher Education*, *63*, 47-57. doi: 10.1016/j.tate.2016.12.009

Kesidou, S., & Roseman, J. E. (2002). How well do middle school science programs measure up? Findings from Project 2061's curriculum review. *Journal of Research in Science Teaching*, *39*, 522-549.

King-Sears, M. E., Brawand, A. E., Jenkins, M. C., & Preston-Smith, S. (2014). Co-teaching perspectives from secondary science co-teachers and their students with disabilities. *Journal of Science Teacher Education*, *25*, 651-680.

Kraft, M. A., Blazar, D., & Hogan, D. (2017). *The effect of teaching coaching on instruction and achievement: A meta-analysis of the causal evidence* (Working paper). Brown University.

Maroney, S. A., Finson, K. D., Beaver, J. B., & Jensen, M. M. (2003). Preparing for successful inquiry in inclusive science classrooms. *TEACHING Exceptional Children*, *36*(1), 18-25.

Mastropieri, M. A., & Scruggs, T. E. (1992). Science for students with disabilities. *Review of Educational Research*, *62*, 377-411.

Mastropieri, M. A., Scruggs, T. E., & Magnusen, M. (1999). Activities-oriented science instruction for students with disabilities. *Learning Disability Quarterly*, *22*, 240-249.

Mastropieri, M. A., Scruggs, T. E., Mantzicopoulous, P., Sturgeon, A., Goodwin, L., & Chung, S. (1998). "A place where living things affect and depend on each other": Qualitative and quantitative outcomes associated with inclusive science teaching. *Science Education*, *82*, 163-179.

Mastropieri, M. A., Scruggs, T. E., Norland, J. J., McDuffie, K., Tornquist, E. H., & Connors, N. (2006). Differentiated curriculum enhancement in inclusive middle school science: Effects on classroom and high stakes tests. *Journal of Special Education*, *40*, 130-137.

Mayer, R. E. (2009). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed.). New York, NY: Cambridge University Press.

McGrath, A. L., & Hughes, M. T. (2017). Students with learning disabilities in inquiry-based science classrooms: A cross-case analysis. *Learning Disability Quarterly*. Advance online publication. doi:10.1177/0731948717736007

McKeown, M. G., & Beck, I. L. (2004). Direct and rich vocabulary instruction. In J. F. Baumann, & E. J. Kame'enui (Eds.), *Vocabulary instruction: Research to practice* (pp. 13-27). New York, NY: Guilford Press.

Moin, L. J., Mageira, K., & Zigmond, N. (2009). Instructional activities and group work in the US inclusive high school co-taught science class. *International Journal of Science and Mathematics Education*, *7*, 677-697.

Mutch-Jones, K., Puttick, G., & Minner, D. (2012). Lesson study for accessible science: Building expertise to improve practice in inclusive science classrooms. *Journal of Research in Science Teaching*, *49*, 1012-1034.

Nagy, W. (2005). Why vocabulary instruction needs to be long-term and comprehensive. In E. H. Hiebert, & M. L. Kamil (Eds.), *Teaching and learning vocabulary: Bringing research to practice* (pp. 27-44). Mahwah, NJ: Lawrence Erlbaum.

National Center for Education Statistics. (2015). *The nation's report card: Science 2015* (NCES 2011-451). Washington, DC: Institute of Education Sciences, U.S. Department of Education.

National Science Teachers Association. (2017). *NSTA position statement: Students with exceptionalities*. Retrieved from http://static.nsta.org/pdfs/PositionStatement_Exceptionalities.pdf

Palincsar, A. S., Magnusson, S. J., Cutter, J., & Vincent, M. (2002). OSEP research institutes: Bridging research and practice: Supporting guided-inquiry instruction. *TEACHING Exceptional Children*, *34*(3), 88.

Quinn, H., Schweingruber, H., & Keller, T. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: The National Academies Press.

Rizzo, K. L., & Taylor, J. C. (2016). Effects of inquiry-based instruction on science achievement for students with disabilities: An analysis of the literature. *Journal of Science Education for Students with Disabilities*, *19*(1), 1-16.

Robinson, S. (2002). Teaching high school students with learning and emotional disabilities in inclusion science classrooms: A case study of four teachers' beliefs and practices. *Journal of Science Teacher Education*, *13*, 13-26.

Sáenz, L. M., & Fuchs, L. S. (2002). Examining the reading difficulty of secondary students with learning disabilities: Expository versus narrative text. *Remedial and Special Education*, *23*, 31-41. doi:10.1177/074193250202300105

Schmidt, W. H., Wang, H. C., & McKnight, C. C. (2005). Curriculum coherence: An examination of US mathematics and science content standards from an international perspective. *Journal of Curriculum Studies*, *37*, 525-559.

Scruggs, T. E., Mastropieri, M. A., & Boon, R. (1998). Science education for students with disabilities: A review of recent research. *Studies in Science Education*, *32*, 21-44.

Scruggs, T. E., Mastropieri, M. A., & Okolo, C. M. (2008). Science and social studies for students with disabilities. *Focus on Exceptional Children*, *41*(2), 1-24.

Stahl, S. A. (2005). Four problems with teaching word meanings. In E. H. Hiebert, & M. L. Kamil (Eds.), *Teaching and learning vocabulary: Bringing research to practice* (pp. 95-114). Mahwah, NJ: Lawrence Erlbaum.

Therrien, W. J., Taylor, J. C., Hosp, J. L., Kaldenburg, E. R., & Gorsh, J. (2011). Science instruction for students with learning disabilities: A meta-analysis. *Learning Disabilities Research & Practice*, *26*, 188-203.

U.S. Department of Education. National Center for Education Statistics. (2016). *Digest of education statistics 2015* (NCES 2016-014). Washington, DC: Institute of Education Sciences, U.S. Department of Education.

U.S. Department of Education. National Center for Education Statistics. (2016). *The Digest of Education Statistics* 2014 (NCES 2016-006), Table 204.60. Retrieved from https://nces.ed.gov/fastfacts/display.asp?id=59

Turner, S. (2008). School science and its controversies: Or, whatever happened to scientific literacy. *Public Understanding of Science*, *17*, 55-72.

Vannest, K. J., Mason, B. A., Brown, L., Dyer, N., Maney, S., & Adiguzel, T. (2009). Instructional settings in science for students with disabilities: Implications for teacher education. *Journal of Science Teacher Education*, *20*, 353-363. doi:10.1007/s10972-009-9135-x

Vaughn, S., & Linan-Thompson, S. (2003). What is special about special education for students with learning disabilities? *The Journal of Special Education*, *37*, 140-147.

Villanueva, M. G. (2010). *Integrated teaching strategies model for improved scientific literacy in second-language learners* (Unpublished doctoral thesis). Nelson Mandela Metropolitan University, Port Elizabeth, South Africa.

Wei, R. C., Darling-Hammond, L., & Adamson, F. (2010). *Professional development in the United States: Trends and challenges* (Vol. 28). Dallas, TX: National Staff Development Council.

Yoon, K. S., Duncan, T., Lee, S. W.-Y., Scarloss, B., & Shapley, K. L. (2007). *Reviewing the evidence on how teacher professional development affects student achievement* (Issues & Answers Report, REL 2007-No.033). Washington, DC: Regional Educational Laboratory Southwest, National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Yore, L. D., Hand, B. M., & Florence, M. K. (2004). Scientists' views of science, models of writing, and science writing practices. *Journal of Research in Science Teaching*, *41*, 338-369. doi:10.1002/tea.20008

Author Biographies

Michael J. Kennedy is an Associate Professor of Special Education at The University of Virginia. His research interests include the intersection of professional development, multimedia, and inclusive literacy instruction in content area classrooms.

Wendy J. Rodgers is an Assistant Professor of Special Education at the University of Nevada, Las Vegas. Her research interests include co-teaching, measurement of instruction, and professional development.

John Elwood Romig is a Doctoral Candidate at the University of Virginia. His research interests include writing instruction, curriculum-based measurement, and professional development.

Hannah Morris Mathews is a Doctoral Candidate at the University of Virginia. Her research interests include teacher preparation, literacy instruction, and professional development.

Katherine N. Peeples is a Doctoral Student at the University of Virginia. Her research interests include the design and delivery of feedback for pre- and inservice teachers