# Teachers' Pedagogical Content Knowledge in Mathematics and Science

A Cross-Disciplinary Synthesis of Recent DRK-12 Projects

David I. Miller, Isabella Pinerua, Jonathan Margolin, and Dean Gerdeman

**April 2022**

American Institutes for Research®  |  AIR.ORG

AIR®

Advancing Evidence.
Improving Lives.

# Teachers' Pedagogical Content Knowledge in Mathematics and Science

## A Cross-Disciplinary Synthesis of Recent DRK-12 Projects

David I. Miller, Isabella Pinerua, Jonathan Margolin, and Dean Gerdeman

**APRIL 2022**

# Acknowledgements

# Contents

# Figures

# Tables

# Project Spotlights

# Executive Summary

## Context and Our Focus

Teachers' pedagogical content knowledge (PCK) is a complex, multifaceted construct that is widely seen as foundational to the act of teaching. In this synthesis, we investigated how the National Science Foundation's (NSF's) recent research investments have advanced understanding and supported the development of teachers' PCK in PK–12 mathematics and science education. In the 5 years from 2011 to 2015, NSF's Discovery Research PK–12 program (DRK-12) funded or cofunded 27 projects relating to PCK, totaling $62 million awarded.

## Findings

These 27 DRK-12 projects primarily applied correlational/observational and longitudinal methods (rather than quasi-experimental or experimental methods), often targeting teaching in the middle school grades. Our synthesis of empirical findings focused on how these projects studied PCK, including its measurement, development, and relationship to teaching and student learning.

- **Measurement of PCK**. One major cross-cutting contribution was methodological advances in the approaches to measure PCK, including quantitative tests (e.g., with multiple-choice answers), qualitative interviews, lesson observations, and clinical simulations. The projects contrasted the affordances and challenges of these methods, highlighting how selecting a specific measurement approach requires aligning the research goal with the method's affordances.

- **Development of PCK**. Many projects also studied the development of PCK, especially through professional development (PD) programs for in-service teachers. Most intervention studies were early-phase investigations, focused on designing and refining new PD programs. The studies highlighted some common design principles for fostering PCK, such as (a) analyzing student work, (b) engaging teachers in active learning, and (c) situating the PD in classroom contexts.

- **Relationship to Teaching Practice and Student Learning**. Compared to projects characterizing and improving teachers' PCK, far fewer projects investigated the relationship between teachers' PCK with teaching practice and student learning. However, the projects that investigated these relationships indicated the value of doing so. One key message was that PD intervention developers must be intentional in connecting lessons about PCK to concrete plans for improving classroom practice.

## Implications for the DRK-12 Portfolio

Our discussion of these projects' findings considers opportunities and priorities for future research, critically examining potential gaps in the DRK-12 portfolio. The recent advancements in measurement approaches make PCK a topic ripe for further exploration, and intervention studies produced promising proofs of concept for developing teachers' PCK. More critically, however, limitations in the evaluation designs produced limited causal evidence on what interventions can effectively improve teachers' PCK. Also, comprehensive understanding remained elusive about the complex relationships of PCK with teaching practice and student learning.

These observations suggest both limitations in the funded DRK-12 portfolio and key opportunities for future projects. The time may be particularly ripe for more impact studies on promising interventions for enhancing teachers' PCK, given the prior foundational design research and advances in PCK measurement approaches. Limitations in the reviewed DRK-12 portfolio also suggest opportunities for studying how PCK fits into a complex system of broader teacher professional knowledge, teacher practice, and student learning. Pursuing such methodological, theoretical, and empirical advances about PCK could help support DRK-12's mission to significantly enhance PK–12 science, technology, engineering, and mathematics teaching and learning.

# Why This Topic?

Effective teacher training for PK–12 science, technology, engineering, and mathematics (STEM) education is crucial for creating supportive learning environments and, ultimately, preparing students to later succeed in STEM-related college majors and careers (Carnegie–Institute for Advanced Study Commission on Mathematics and Science Education, 2009). For decades, educational researchers have argued that professional development (PD) opportunities in these fields should build on and develop teachers' *pedagogical content knowledge* (PCK; Ball & Cohen, 1999; Desimone, 2009).

Shulman (1986) coined the term PCK in his landmark paper, distinguishing it from both general content knowledge and general pedagogical knowledge. For instance, teachers could understand fractions well (content knowledge) yet lack awareness of common student ideas about the topic (an example of mathematical PCK). Knowledge of how to teach fractions is also distinct from knowledge of broad, domain-general pedagogical strategies or classroom management techniques (pedagogical knowledge). Shulman (1987) characterized PCK as "that special amalgam of content and pedagogy that is uniquely the province of teachers, their own special form of professional understanding" (p. 8).

Although scholars' exact definitions vary, the broad construct of PCK has guided decades of conducting educational research studies and designing PD programs (Gess-Newsome, 2015). For instance, this construct has helped to explain the typically low correlations between teacher content knowledge and student outcomes, highlighting that integrating content and pedagogy is critical for supporting student learning (Wayne & Youngs, 2003). Thus, PCK has appealed to researchers across many subject areas but particularly in mathematics and science education.

This report presents our cross-disciplinary synthesis on teachers' PCK in mathematics and science education, as studied in recent projects funded by the National Science Foundation's (NSF's) Discovery Research PK–12 (DRK-12) program. We identified PCK as one key area of NSF investment in PK–12 STEM education, based on a structured review of recent DRK-12 award abstracts.[1] We found that, in the 5 years spanning 2011 to 2015, the DRK-12 program funded or cofunded 27 projects relating to PCK, totaling $62 million awarded.

> In the 5 years spanning 2011 to 2015, the DRK-12 program funded or cofunded 27 projects relating to PCK, totaling $62 million awarded.

---

[1] This report's synthesis on NSF-funded PCK research comes from a broader NSF project, *Advancing Methods and Synthesizing Research in STEM Education* (DRL-1813777), which aims to synthesize evidence of innovation and discovery in recent DRK-12 projects.

Teachers' Pedagogical Content Knowledge in Mathematics and Science

We synthesized these 27 recent PCK-related projects to understand recent NSF-funded innovations in methods, theory, and empirical understanding of PCK. We first briefly review relevant theoretical perspectives that helped us examine how researchers conceptualized PCK differently across projects and educational disciplines.

## Varying Definitions of PCK

Although PCK is widely seen as foundational to the act of teaching, researchers disagree on how to define the construct (Chan & Hume, 2019; Depaepe et al., 2013; Fernandez, 2014). For instance, controversy remains about the grain size of PCK (Carlson et al., 2019), such as whether it exists at the level of the broad discipline (e.g., mathematics), the topic (e.g., basic arithmetic), or the concept (e.g., equivalence in addition problems). In her review of models of PCK, Fernandez (2014) noted that researchers often speak of PCK as if it were clearly defined, without providing their own definition or even being explicit about the theoretical model they are using.

The following sections briefly review the theoretical models commonly used in two disciplinary fields, mathematics education and science education, in which PCK has often been studied. Although these two fields have unique disciplinary traditions for studying PCK, they have also wrestled with similar underlying conceptual challenges (e.g., understanding the context specificity of PCK).

## Defining PCK in Mathematics Education

In the field of mathematics education, PCK is often studied within the context of a broader concept of professional knowledge known as *mathematical knowledge for teaching* (MKT; Ball et al., 2008). Ball et al. motivated their MKT framework by considering the fundamental question: What do teachers need to know to teach mathematics effectively? This practice-based question motivated the construct of MKT, which includes both PCK and subject matter knowledge (see Figure 1).

## Figure 1. Domains of Mathematical Knowledge for Teaching

**SUBJECT MATTER KNOWLEDGE**  **PEDAGOGICAL CONTENT KNOWLEDGE**

Common content knowledge (CCK)

Specialized content knowledge (SCK)

Horizon content knowledge

Knowledge of content and students (KCS)

Knowledge of content and teaching (KCT)

Knowledge of content and curriculum
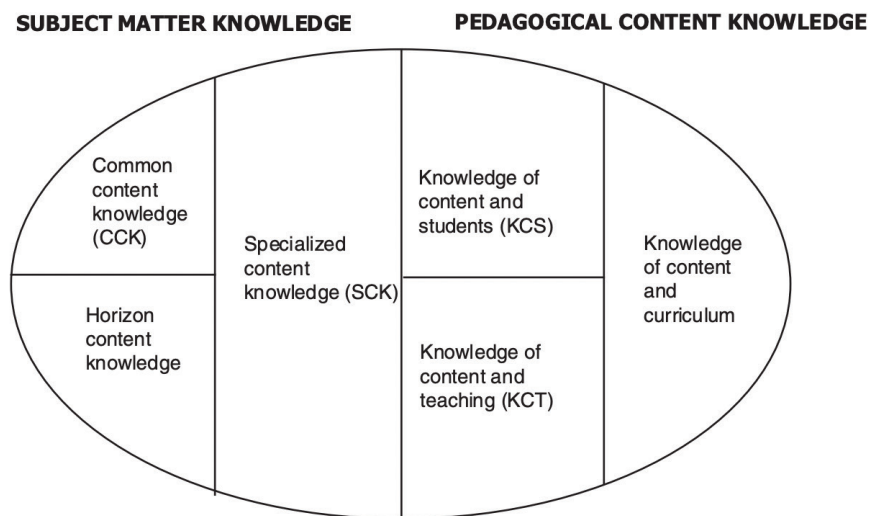
*Note.* Adapted from Ball et al. (2008).

Ball et al.'s (2008) framework divided PCK into three components regarding content-specific knowledge about (a) student thinking, (b) instructional strategies, and (c) curriculum (the right half of Figure 1). The model elaborated on the first two PCK components, but it tentatively included the third (i.e., PCK about curriculum), questioning if curricular knowledge should be a separate knowledge category or should be part of the PCK category about instructional strategies.

The MKT framework was broader than PCK by including the subject matter knowledge needed to teach effectively (the left half of Figure 1). Although distinct from PCK, this content knowledge can still be specific to the mathematical demands of teaching (e.g., knowing how to explain why you can "add a zero" when multiplying by 10; Ball et al., 2008, p. 400).

Depaepe et al. (2013) noted three key merits of the concept of MKT, drawing from their systematic review of the relevant mathematics education literature:

- MKT is built on empirical findings about how teachers apply knowledge in classroom practice.

- The construct was specified well enough to enable researchers to operationalize it through an assessment, the MKT test (Hill et al., 2004; Hill et al., 2005).

- Research with this assessment has illustrated how MKT can shape student learning.

Although the MKT framework has been a central foundation, researchers disagree about how to best conceptualize and measure these knowledge constructs (see Copur-Gencturk &

Lubienski, 2013; Kaarstein, 2014). For instance, Petrou and Goulding (2011) argued that the MKT model's distinction between PCK and specialized content knowledge has not been well defined or empirically supported. Scholars also disagree whether teachers' PCK is static and can be assessed outside the classroom (e.g., on a quantitative MKT test) or whether it is meaningful only in the context of its application (e.g., requiring classroom observations to observe PCK in action; Rowland & Ruthven, 2011).

Despite these disagreements, Depaepe et al.'s (2013) review found some key similarities in various theoretical models of mathematical PCK. For instance, researchers broadly agree that PCK "connects content and pedagogy . . . is specific to teaching particular subject matter, and [that] content knowledge is an important and necessary prerequisite [for developing PCK]" (p. 22). PCK is also widely regarded as foundational to effective teaching in mathematics, even if its exact definition varies.

## Defining PCK in Science Education

Akin to mathematics education, considerable debate remains on how to conceptualize PCK in science education (Abell, 2008). For instance, Fernandez (2014) reviewed 10 models of science teachers' PCK, each having its own graphical representation of how PCK relates to and contrasts with other professional knowledge bases. Gess-Newsome (2015) noted that some of the controversies include the following:

- What is the grain size of PCK (e.g., discipline, topic, concept)?

- Is PCK a knowledge base or a skill, or both?

- Is PCK knowledge held by the community or is it a teacher attribute?

- Can PCK be measured separately from the act of teaching?

These questions resemble similar scholarly discussions in the mathematics education community (e.g., such as whether PCK can be measured separately from the act of teaching; see Rowland & Ruthven, 2011, for a discussion of this consideration in mathematics education).

In response to these concerns, an expert group of 22 PCK science education researchers conducted a summit in 2012, resulting in a model of science teacher professional knowledge and skill, now called the *consensus model* (Gess-Newsome, 2015). This model defined PCK as the "knowledge of, reasoning behind, and planning for teaching a particular topic in a particular way for a particular purpose to particular students for enhanced student outcomes" (Gess-Newsome, 2015, p. 36). A second PCK summit in 2016 led to the *refined consensus model* (Carson et al., 2019), which further elaborated on the PCK construct, placing it within a complex system of professional knowledge and experiences that inform science teachers' practice (see Figure 2).

Teachers' Pedagogical Content Knowledge in Mathematics and Science

**Figure 2. Refined Consensus Model of Science Teachers' PCK**
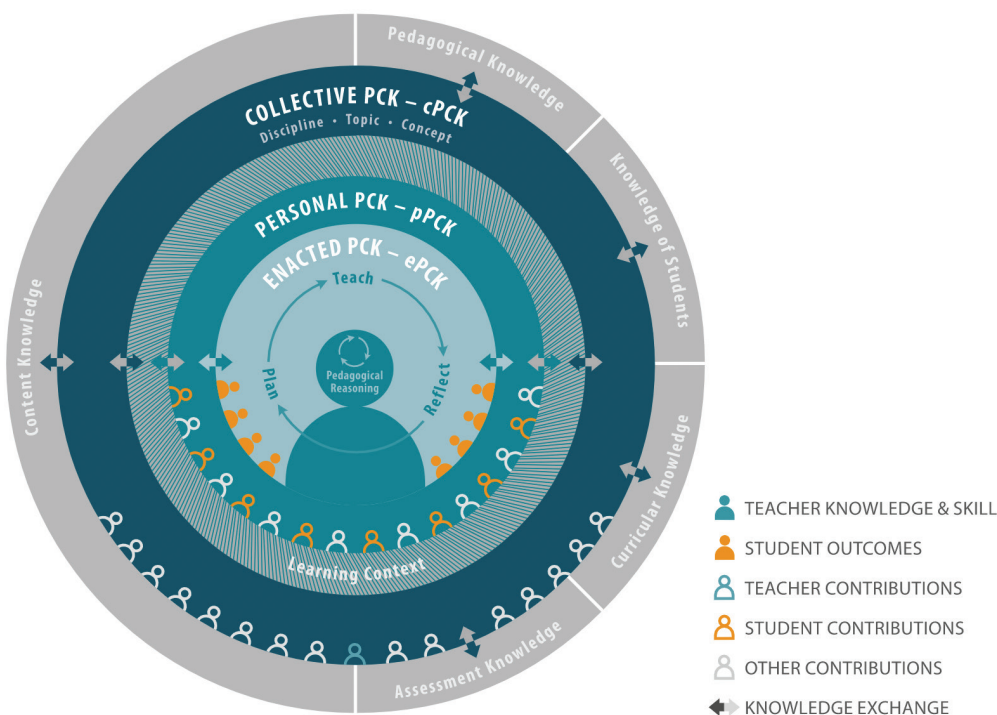


*Note.* Image sourced from Carson et al. (2019).

The developers of these consensus models aimed to address controversies in the field by distinguishing different kinds of PCK. For instance, the refined model detailed three realms of science PCK: (a) enacted PCK (i.e., the knowledge and skills used when teaching specific content in practice), (b) personal PCK (i.e., a teacher's cumulative reservoir of knowledge about teaching specific content), and (c) collective PCK (i.e., an amalgam of multiple teachers' knowledge about teaching specific content). Questions such as "Can PCK be measured separately from the act of teaching?" can be answered as "Yes" by noting the distinction between enacted versus personal PCK. Likewise, PCK can exist at the concept, topic, and even discipline levels, although key distinctions exist between these different PCK grain sizes.

## Defining PCK for Our Synthesis

As reviewed previously, educational researchers vary in how they characterize PCK, its components, and its connection to broader professional teacher knowledge categories. Despite studying PCK in distinct scholarly communities, mathematics and science education researchers have wrestled with analogous conceptual challenges about characterizing this multifaceted construct.

Our synthesis did not aim to "solve" these definitional issues. Rather, we aimed to understand how this conceptual context shaped recent empirical studies funded by NSF's DRK-12 program.

Teachers' Pedagogical Content Knowledge in Mathematics and Science

We therefore did not adopt a specific theoretical model of PCK because we wanted to understand the conceptual complexity of PCK as used across empirical DRK-12 projects. However, we heavily drew from the coding structure in Depeape et al.'s (2013) systematic review, which aimed to characterize differences between PCK models and the studies using them. To bound our review, we included DRK-12 projects that studied at least one of the four (presumed) PCK components regarding teacher knowledge about (a) student thinking, (b) instructional strategies, (c) curriculum, or (d) assessment.

We included studies that used a content knowledge for teaching framework (e.g., MKT) because they encompassed PCK, although we excluded studies purely about teachers' content knowledge (without also studying PCK). We included studies about teaching practice (e.g., using observation protocols) only if they included a specific focus on PCK in their theoretical framing (e.g., studying PCK in action).

## Improving Teaching and Learning in DRK-12 Projects

Beyond synthesizing efforts to characterize and measure teachers' PCK, we investigated how researchers used understanding about PCK to advance DRK-12's central mission: enhance STEM learning and teaching in the PK–12 grade levels. We therefore also reviewed DRK-12 projects that studied how PCK develops (e.g., through PD programs) and how it relates to teaching practice and student learning.

Teacher PCK is important, in part, because this knowledge can enable teachers to create more supportive learning environments (e.g., better respond to students, develop more effective lesson plans; Baumert et al., 2010). These supportive learning environments could then improve student outcomes. For instance, empirical studies have found that teachers' PCK correlates with instructional quality and student learning gains, even after controlling for teachers' content knowledge (e.g., Baumert et al., 2010; Griffin et al., 2009). Several studies have also found that PD interventions can improve teacher PCK and student outcomes (e.g., Heller et al., 2012; Roth et al., 2011; Seymour & Lehrer, 2006; Taylor et al., 2017; Tirosh et al., 2011).

We structured our review of empirical findings based on four key lines of empirical research about PCK: (a) measurement of PCK, (b) development of PCK, (c) relationship to teaching practice, and (d) relationship to student learning (Depaepe et al., 2013). Before detailing specific findings, we first present descriptive information about the recent DRK-12 projects we included and what aspects of PCK they studied.

# What Was Studied?

Although PCK has often been studied in separate disciplines, some cross-cutting considerations apply across traditional research boundaries, such as mathematics and science education research. Our synthesis studied how the multidisciplinary community of DRK-12-funded researchers conceptualized and investigated the construct of PCK in PK–12 mathematics and science education research. We focused on recent DRK-12 projects to identify new advances in understanding and characterize emerging best practices for designing effective PD. These set of DRK12 projects represented a substantial investment in research related to PCK in STEM education, though they do not necessarily reflect the broader recent research in this area.

## Our Synthesis Approach

Appendix A details our review methodology, which we briefly summarize here. We examined DRK-12 projects with an original award date spanning January 2011 to December 2015, to focus on recently completed or close-to-completion projects. We downloaded the award abstracts for all DRK-12 awards in this date range using NSF's website. When screening projects for their relevance to PCK, we defined PCK as the knowledge of, reasoning behind, and planning for teaching a specific educational topic or domain (Gess-Newsome et al., 2015, p. 36). Some award abstracts mentioned the term "pedagogical content knowledge" explicitly, whereas others described the concept in other ways (e.g., "teachers' knowledge of students' thinking"), yielding 27 eligible projects.

**Figure 3. Overview of Our Synthesis of 27 DRK-12 Projects Related to PCK**

| DRK-12 Projects With Eligible Award Date **376 Projects** | → | Review Abstracts to Identify PCK-Related Projects **27 Projects** | → | Search/Screen for Products From Each Project **50[a] Products** | → | Code and Analyze Project Findings |

[a] More than 50 products were identified, but for feasibility, we restricted our synthesis to the one to three products per project that were most directly related to PCK and most completely reported (e.g., favoring peer-reviewed journal articles rather than conference presentations or short project summaries).

We then systematically looked for products that these 27 projects generated, searching six sources: Web of Science, ERIC, PyscINFO, Google Scholar, Research.gov, and the Community for Advancing Discovery Research in Education (CADRE) website. For instance, we used Google Scholar to identify documents whose full text contained the numeric award ID and the phrases words "NSF" or "National Science Foundation." We also sent emails to all project principal investigators (PIs) asking for additional products that our searches may have missed. For each

Teachers' Pedagogical Content Knowledge in Mathematics and Science

NSF project, we identified between one to three products that were most closely related to PCK, prioritizing the inclusion of more complete, peer-reviewed reports (e.g., journal articles) rather than less complete, less final reports (e.g., conference presentations). This process yielded 50 study reports (i.e., products) that formed our synthesis of the 27 projects' findings (see Figure 3).

We coded and analyzed the products using two approaches: (a) structured coding based on a priori categories (e.g., component of PCK studied) and (b) qualitative narrative review of project findings (supported by the software NVivo). The following section details results from the first approach, which was informed by coding protocols from prior systematic reviews on PCK (Depaepe et al., 2013; Evens et al., 2015; Hoover et al., 2016; Schneider & Plasman, 2011). Later sections present the second approach.

## What PCK-Related Problems or Topics Were Studied?

We coded how many of the 27 reviewed NSF projects addressed the four following major lines of empirical research on PCK (adopting categories from Depaepe et al., 2013):

- **Nature of PCK**. Conceptualizing, measuring, and characterizing teachers' PCK (21 projects)

- **Development of PCK**. Investigating the development of teachers' PCK, including studies of both interventions and natural acquisition of PCK across time (14 projects)

- **Relationship to Teaching Practice**. Examining the relationship between teachers' PCK and teaching quality or instructional practices (5 projects)

- **Relationship to Student Learning**. Examining the impact of teachers' PCK on student learning (5 projects)

These counts are not mutually exclusive. For instance, a project could conduct an in-depth characterization of teachers' PCK (yielding a count for the "nature of PCK" category) and also study how a PD program improves PCK (yielding a count for the "development of PCK" category). We conducted the coding at the product level and then summarized frequencies at the project level (e.g., 14 NSF projects had at least one product about the development of PCK).

The most frequently studied lines of research were on the nature and development of PCK, with 21 and 14 NSF projects, respectively. In contrast, the two lines of research on relationships to teaching practice and student learning were studied less often (in five projects each).

The top row in Figure 4 shows these counts graphically, giving darker colors to major lines of PCK research studied more frequently. Other rows of Figure 4 show frequency counts for the PCK component and the major disciplinary field studied.

This figure also shows cross tabulations between the major line of PCK research and two other project characteristics (PCK component and disciplinary field studied). For instance, the PCK component of teachers' knowledge about instructional strategies was studied in all six projects that examined the relationship between PCK and teaching practice. This figure disaggregates counts by the major line of PCK research (i.e., the vertical columns) because the major line of PCK research was centrally important in the organization of our review's later sections about synthesizing specific empirical findings. However, relative frequencies in Figure 4 were generally similar across major lines of research.

**Figure 4. Number of NSF Projects by Major Line of PCK Research (Vertical Columns) and PCK Component and Major Disciplinary Field (Horizontal Rows)**

| | Any line of research | Nature of PCK | Development of PCK | Relationship to teaching practice | Relationship to student learning |
|---|---|---|---|---|---|
| **Total number of projects** | 27 | 21 | 14 | 5 | 5 |
| **Component of PCK studied** | | | | | |
| Knowledge of students' thinking | 23 | 19 | 13 | 4 | 5 |
| Knowledge of instructional strategies | 18 | 15 | 9 | 5 | 3 |
| Knowledge of curriculum | 4 | 1 | 3 | 1 | 1 |
| Knowledge of assessment | 2 | 2 | 1 | 0 | 0 |
| **Major disciplinary field** | | | | | |
| Mathematics | 15 | 12 | 9 | 3 | 3 |
| Science | 14 | 11 | 7 | 4 | 3 |
| Engineering | 3 | 2 | 1 | 1 | 0 |

The most frequently studied PCK components[2] were teachers' knowledge of students thinking and instructional strategies (23 and 18 NSF projects, respectively), which are the two components that Shulman (1986) featured in his original conceptualization of PCK, whereas fewer projects studied teachers' knowledge of curriculum and assessment strategies (four and two projects, respectively). These projects were roughly evenly split between mathematics and science education. The most common mathematical topic areas were fractions (five projects), general mathematics/not specified (five projects), and basic arithmetic (four

---

[2] We constrained our review to four (presumed) PCK components: teacher knowledge about (a) students' thinking, (b) instructional strategies, (c) curriculum, and (d) assessment. In doing so, we do not imply that these are the definitive components of PCK; rather, this categorization served as a practical tool for structuring our review.

Teachers' Pedagogical Content Knowledge in Mathematics and Science

projects). The common science topic areas were general science/not specified (five projects), biology (four projects), and physical science (three projects). In addition, three projects examined PCK in engineering as an emerging area of research.

## What Research Methods Were Used?

As shown in Figure 5, the reviewed DRK-12 projects usually used correlational/ observational methods (14 projects) and longitudinal methods (15 projects). Quasi-experimental methods (six projects) and experimental methods (three projects) were used less often.

For research on the development of PCK, longitudinal methods were clearly the dominant method (12 of 14 projects in the PCK development research category), often reflecting single group pre-post designs, such as measuring teachers' PCK before and after a PD intervention.

**Figure 5. Number of NSF Projects by Major Line of PCK Research (Vertical Columns) and Research Methods (Horizontal Rows)**

| Study design | Any line of research | Nature of PCK | Development of PCK | Relationship to teaching practice | Relationship to student learning |
|---|---|---|---|---|---|
| Correlational | 14 | 13 | 3 | 4 | 2 |
| Longitudinal | 15 | 10 | 12 | 3 | 5 |
| Quasi-experimental | 6 | 2 | 4 | 2 | 3 |
| Experimental | 3 | 2 | 3 | 0 | 1 |

| PCK measurement method | Any line of research | Nature of PCK | Development of PCK | Relationship to teaching practice | Relationship to student learning |
|---|---|---|---|---|---|
| Quantitative test/survey | 15 | 10 | 12 | 3 | 4 |
| Qualitative interview | 15 | 15 | 6 | 3 | 2 |
| Lesson observations/simulation | 11 | 11 | 6 | 4 | 5 |
| Meeting observation | 4 | 2 | 1 | 2 | 1 |
| Document analysis | 4 | 3 | 1 | 1 | 1 |

The most common approaches for measuring PCK were quantitative tests/surveys (15 projects), qualitative interviews (15 projects), and lesson observations/clinical simulations (12 projects). Less frequently, projects used meeting observations (e.g., at a PD workshop) or document analysis (e.g., review of teachers' written lesson plans for evidence of PCK). Quantitative tests and surveys were especially common for projects studying the development of PCK.

Sample sizes were generally small, with a median number of 21 teachers per study. One quarter of the studies had 12 teachers or fewer, and another quarter had 49 teachers or more.

Teachers' Pedagogical Content Knowledge in Mathematics and Science

## What Teacher Populations Were Studied?

Most projects (23 of 27 projects) studied in-service teachers, although some (eight projects) included preservice teachers (see Figure 6). The most typical grade level taught was middle school (15 projects), followed by high school (10 projects) and upper elementary (eight projects). Notably, no project studied teachers who taught (or planned to teach) early elementary grade levels (grades K–2) or prekindergarten. Relative frequencies did not substantially vary by the major line of PCK research (Figure 6).

**Figure 6. Number of NSF Projects by Major Line of PCK Research (Vertical Columns) and Teacher Population (Horizontal Rows)**

| | Any line of research | Nature of PCK | Development of PCK | Relationship to teaching practice | Relationship to student learning |
|---|---|---|---|---|---|
| **Teacher experience** | | | | | |
| In-service teachers | 23 | 17 | 13 | 5 | 5 |
| Preservice teachers | 8 | 7 | 3 | 3 | 2 |
| **Teaching grade level** | | | | | |
| Early elementary (Grades K–2) | 0 | 0 | 0 | 0 | 0 |
| Late elementary (Grades 3–5) | 8 | 7 | 2 | 1 | 1 |
| Middle school (Grades 6–8) | 15 | 12 | 7 | 3 | 1 |
| High school (Grades 9–12) | 10 | 6 | 4 | 2 | 1 |

## What Interventions or Resources Were Developed?

The most common type of developed resource was PD workshop materials, such as videos to analyze teaching practices and student learning, lesson plans and discussion prompts for summer PD programs, and written examples of student thinking (17 projects). Also, five projects developed preservice course materials for fostering PCK. Lastly, 10 projects developed PCK assessments, although other projects also reviewed the measurement properties of existing PCK assessments, as reviewed in the following section.

Teachers' Pedagogical Content Knowledge in Mathematics and Science

# What Was Learned About the Measurement of PCK?

One key area of progress for the field was DRK-12 projects that developed and refined approaches for measuring teachers' PCK. These approaches fell into three broad categories: (a) quantitative tests and surveys, (b) qualitative interviews, and (c) lesson observations and clinical simulations. The following summary contrasts the affordances, challenges, and best practices for each measurement approach, as reported in the products we reviewed (see Table 1). Readers interested in specific disciplines or topics (e.g., what teachers know about students' algebra ideas) may consult Table B1 in Appendix B for more detailed summaries of project findings related to characterizing teachers' PCK.

**Table 1. Summary of DRK-12 Researchers' Considerations for Using Three Common PCK Measurement Approaches**

| Approach | Affordances | Challenges | Best practices |
|---|---|---|---|
| Quantitative tests and surveys | • Can deliver at scale<br>• Supports statistical analysis<br>• Accommodates multiple-choice and constructed-response formats | • Teacher reasoning depends heavily on pedagogical context<br>• Limited opportunities for teachers to explain their reasoning<br>• Multiple answers could be defensible depending on context | • Consult expert teachers for initial item review<br>• Use qualitative data to refine items<br>• Carefully consider the design of the items' pedagogical context |
| Qualitative interviews | • Allows teachers to explain their reasoning in detail<br>• Can help reveal connections between different types of PCK | • Teachers may have tacit PCK that is hard to articulate<br>• Difficult to administer at large scales | • Develop interview protocols before conducting interviews<br>• Train coders how to systematically code and analyze data<br>• Triangulate with other data sources |
| Lesson observations and clinical simulations | • Lesson observations offer real-world authenticity to show PCK in action<br>• Clinical simulations standardize the evaluation context | • Situation-specific demands make real-world contexts difficult to interpret<br>• Simulated scenarios may oversimplify the PCK needed for real-world teaching tasks | • Specify criteria for selecting specific teacher-student interactions<br>• Develop concrete indicators of evidence for teachers using their PCK in action |

Teachers' Pedagogical Content Knowledge in Mathematics and Science

## Quantitative Tests and Surveys

Four DRK-12 projects developed novel quantitative tests and surveys related to PCK, covering topics such as pedagogical beliefs about scientific argumentation (McNeill et al., 2014; Katsh-Singer et al., 2016); content knowledge for teaching energy in science (Etkina et al., 2018); mathematical knowledge for teaching with visual representations (DePiper & Driscoll, 2018; Louie & Nikula, 2019); and self-efficacy for having the PCK to teach robotics-based lessons (Rahman et al., 2017). One other project used qualitative think-aloud interviews to better understand existing quantitative measures in the mathematics domain (Hoover & Lai, 2017).

### *Approach, Affordances, and Challenges*

The quantitative PCK measures often featured vignettes or specific instructional scenarios, asking teachers to evaluate students' thinking or select an effective teaching action. For instance, one item presented two students talking about the types of energy involved in dribbling a basketball, as shown in the following spotlight for the project, *Assessing, Validating, and Developing Content Knowledge for Teaching Energy* (Project Spotlight 1). Multiple-choice response formats were common (see Questions 1 and 2 in Project Spotlight 1 as an example). However, researchers sometimes also used constructed-response formats (see Question 3), which were systematically scored using quantitative rubrics.

One key lesson learned about PCK instrument development concerned the benefits and challenges of grounding these vignettes in specific pedagogical contexts. Researchers saw including rich pedagogical context as fundamental to the valid measurement of PCK given the context-specific nature of PCK and effective instruction (Hoover & Lai, 2017; McNeill et al., 2014). However, these projects identified several challenges in designing these items' contexts. For instance, the contextual features may not be rich enough to make nuanced instructional decisions or could even allow for multiple appropriate responses, which is a major concern for multiple-choice response formats (McNeill et al., 2014). Hoover and Lai (2017) conducted think-aloud interviews to understand how expert teachers interpreted the pedagogical contexts for the widely used MKT items.

> One key lesson learned about PCK instrument development concerned the benefits and challenges of using vignettes that grounded assessment items in specific instructional scenarios.

### *Best Practices*

These projects offered several best practice principles for developing PCK assessment items: (a) specify the key theoretical constructs of interest before designing the items, (b) carefully consider the instructional context described, (c) consult experienced teachers and content experts during initial item review, (d) use think-aloud interviews or other qualitative approaches for understanding teachers' thought processes in answering the items, (e) pilot test

Teachers' Pedagogical Content Knowledge in Mathematics and Science

items with the target teacher population, (f) evaluate psychometric metrics, and (g) iteratively refine the items through multiple testing cycles.

# Project Spotlight 1: Measurement of PCK

*Assessing, Validating, and Developing Content Knowledge for Teaching Energy* (NSF awards #1222777, 1222732, 1222580, 1222598; total funded amount = $3 million)

## Why Spotlight This Project?

It demonstrates three distinct approaches (quantitative test, PD meeting observations, lesson observations) for measuring teachers' PCK, using rigorous qualitative and quantitative methods.

## What Was Studied?

The project focused on the pedagogical and disciplinary content knowledge that middle school and high school physical science teachers need to teach energy topics.

## What Was Found?

- **Quantitative Test**. Etkina et al. (2018) developed, piloted, and psychometrically validated a quantitative Content Knowledge for Teaching Energy (CKT-E) instrument, including both multiple-choice and constructed-response item formats (see the example below).
- **Meeting Observations**. Wittmann et al. (2017) studied teachers' discussion about a CKT-E item during a PD session. Teachers demonstrated limited explicit awareness of the common student metaphor of energy as a substance-like quantity.
- **Lesson Observations**. Robertson et al. (2017) developed a qualitative methodology for inferring teachers' CKT-E based on selecting and analyzing classroom episodes that illustrated how teachers inferred, restated, and evaluated students' ideas about energy.

## Example CKT-E item (Etkina et al., 2018)

Two students in Ms. Engel's physics class are discussing the energetics of dribbling a basketball on a wooden floor. They agree that all of the kinetic energy gets converted into elastic energy for an instant when the basketball is compressed the most. They also agree that many objects can be modeled as springs, even basketballs and wooden floors. They are uncertain about whether there would be equal amounts of elastic energy in the ball and the floor. They call Ms. Engel over to share their ideas with her and get some help.

Marcos says, "We were thinking that when the ball compresses against the floor, the forces that the ball and the floor exert on each other would be equal and opposite, so maybe the amount of elastic energy in the floor is the same as the elastic energy in the ball."

Louisa responds, "I get that the forces are the same, but I am thinking that the ball compresses more than the floor, so shouldn't there be more energy stored in the ball?"

Marcos replies, "But the floor is more rigid and would have a higher spring constant. I think the larger k of the floor compensates for the smaller $\Delta x$ in the $\frac{1}{2}k(\Delta x)^2$ equation, and the elastic energies are the same."

1. Is Marcos correct that the elastic energy of the ball and the floor would be the same?
   - A. Yes. (31%)
   - B. No. The elastic energy of the ball would be greater. (51%)
   - C. No. The elastic energy of the floor would be greater. (1%)
   - D. There is not enough information to compare these energies. (17%)

2. Which of the following activities would be most likely to provide Ms. Engel's students with additional insights about the relative amounts of elastic energy during the bounce of the basketball?
   - A. They could measure the spring constants and the displacements of both the floor and the ball and use those to compare the elastic energies. (10%)
   - B. They could compare the elastic energies of two non-identical springs when they are compressed with the same force. (70%)
   - C. They could do an experiment to see if a basketball bounces higher on a soft carpet surface or a hard concrete floor. (14%)
   - D. They could do an experiment to show that the same basketball will not bounce as high off the gym floor if it has first been put in a freezer. (6%)

3. Explain your selection and how the activity you selected might provide the students with additional insights about the relative elastic energy during the bounce of a basketball.

## Qualitative Interviews

DRK-12 projects often used interview-based methods to gather qualitative insights about teachers' PCK. These projects spanned several topics such as the instructional value of scientific argumentation activities (Katsh-Singer et al., 2016), beliefs about teaching English learners in science classrooms (Lyon et al., 2016), the role of using interactive simulations in algebra classes (Findley et al., 2017), knowledge for teaching basic arithmetic and fractions (Hoover & Lai, 2017), student misconceptions about ecosystems and the particle model of matter (Smith et al., 2017; Smith et al., 2018), teachers' explanations of students' mathematical reasoning (Hodkowski, 2018), and student learning trajectories in mathematics (Castro Superfine & Li, 2017).

### *Approach, Affordances, and Challenges*

Some projects detailed the methodological strengths and challenges of using qualitative interviews. For instance, one science education project compared the affordances of open-ended online survey questions versus telephone interviews. Although teachers often gave vague answers to open-ended surveys (e.g., "please describe the ideas or misconceptions your students have . . . about the particulate model of matter"), they usually gave much more detailed interview responses that showed connections between different types of PCK (Smith et al., 2017). Nevertheless, teachers frequently strayed off topic during the interviews, and their responses were often ambiguous to interpret. Also, teachers may have tacit knowledge that is hard to express in an interview-based setting, even if they can use their tacit PCK in applied classroom settings (Katsh-Singer et al., 2016; Smith et al., 2017; Smith et al., 2018).

> Teachers may have tacit knowledge that is hard to express in an interview-based setting, even if they can use their tacit PCK in applied classroom settings.

### *Best Practices*

Best practice suggestions for using interviews included (a) developing detailed interview protocols in advance, (b) including possible follow-up questions in the protocol for vague or ambiguous interview responses, and (c) coding and analyzing responses using two independent raters and detailed codebooks. Smith and colleagues also suggested delivering open-ended surveys prior to the interview; they argue doing so could help the interviewer prepare individually tailored questions that probe the survey responses about PCK in more depth (Smith et al., 2017; Smith et al., 2018).

## Lesson Observations and Clinical Simulations

PCK measurement approaches based on lesson observations and clinical simulations aim to observe teachers' PCK in action (i.e., as enacted when interacting with real or simulated students). DRK-12 researchers used both approaches for

> Lesson observations and clinical simulations aim to observe teachers' PCK in action.

Teachers' Pedagogical Content Knowledge in Mathematics and Science

several topics, such as content knowledge for teaching energy (Robertson et al., 2017), natural selection (Dotger et al., 2018), engineering design (Crismond & Lomask, 2016), geometry (Dotger et al., 2015), fractions (Hodkowski, 2018), and basic arithmetic (Shaughnessy & Boerst, 2018; Shaughnessy et al., 2018).

## *Approach, Affordances, and Challenges*

PCK measurement approaches using lesson observations are often based on analyzing recorded interactions with students in naturalistic classroom contexts. For instance, in the context of teaching about energy in science, Robertson et al. (2017) developed a methodology for selecting relevant classroom episodes of student-teacher interactions, as well as identifying the task of teaching (e.g., choosing an instructional activity, evaluating student ideas) and the energy content knowledge addressed. The researchers inferred teachers' content knowledge for teaching energy based on how teachers (a) inferred students' model of energy from students' expressions, (b) restated students' ideas, (c) chose an instructional activity to address a misunderstanding, and (d) evaluated student ideas.

Other applications combined lesson observations with other artifacts, such as teachers' instructional logs and postlesson reflections, to create indicators of PCK in action (e.g., Crismond & Lomask, 2016). For instance, one indicator of teachers' PCK was adapting the provided curriculum to meet the learning needs of teachers' specific students.

In contrast, PCK measurement approaches based on clinical simulations involve interactions with students in "standardized" students in controlled settings. These students are typically research staff trained to respond in scripted ways, based on categories of teachers' actions. Building on the long history of clinical simulations in medical education, two DRK-12 projects used this approach to capture teachers' skills in diagnosing students' content-specific thinking in STEM, which are skills that heavily depend on teachers' PCK (Dotger et al., 2015; Dotger et al., 2018; Shaughnessy & Boerst, 2018; Shaughnessy et al., 2018).

For instance, in one study, preservice teachers reviewed a student's work on a specific addition problem, after which they had 5 minutes to interact with a standardized student to diagnose common mathematical misconceptions (Shaughnessy & Boerst, 2018). Figure 7 shows example instructions for this standardized student. Researchers developed a checklist of core desired practices, including (a) eliciting the student's process, (b) probing understanding of key mathematical ideas, (c) attending to the student's ideas, and (d) deploying other moves that support learning.

**Figure 7. Example Instructions for Standardized Students in Clinical Simulations**

<u>What the student is thinking:</u>

- the student is using a version of the "column addition method" for solving multi-digit addition problems
- the student can use the same process to solve addition problems with more than two digits, and the student does understand when/why/how to "combine"
- the student is thinking about 623 as 6 tens and 23 ones

<u>General orientation to responses:</u>

- talk about digits in terms of the place value of the column
- use "combining" to refer to trading/carrying/regrouping
- don't make basic facts errors and don't need to show strategies for adding one-digit numbers
- give the least amount of information that is still responsive to the preservice teacher's question
- if a question is confusing, say something like, "What do you mean?"

<u>Specific responses (*a subset of them*):</u>

| *Preservice teacher question* | *Response* |
|---|---|
| So tell me what you did" or "So how did you get 83? | I did 29 + 36 + 18. |
| **What** did you do first? | I added the tens: 2 + 3 + 1 and I got 6. |
| **How** did you get from 623 to 83? | I had to combine the 6 and the 2. |
| **Why** did you need to combine those numbers? | Because they're both tens. |

A subsequent study contrasted the unique affordances and challenges of lesson observations versus clinical simulations (Shaughnessy et al., 2018). Although the lesson observations capture real-world teaching contexts with greater authenticity, the variability of students and situation-specific demands presents challenges for interpreting teachers' PCK. Clinical simulations address this issue with standardized students but at the cost of real-world authenticity. Shaughnessy and Boerst argued that the two approaches are complementary. Another DRK-12 project developed similar clinical simulation procedures for natural selection (Dotger et al., 2018) and geometry topics (Dotger et al., 2015).

## *Best Practices*

For both lesson observations and clinical simulations, common best practices included (a) specifying criteria for which teacher-student interactions should be coded and analyzed, (b) developing concrete indicators for identifying evidence of teachers' PCK in action, and (c) articulating the key tasks of teaching involved. Additional considerations for clinical simulations included developing response protocols for "standardized students" and training research staff to implement them consistently across teachers.

Teachers' Pedagogical Content Knowledge in Mathematics and Science

# What Was Learned About the Development and Impact of Teacher PCK?

The reviewed DRK-12 projects used new or existing measurement approaches to gain substantive insights about PCK. The following review divides these insights into three major lines of empirical research: (a) the development of PCK, (b) the relationship of PCK to teaching practice, and (c) the relationship of PCK to student learning. Although our review categorizes the results into these three lines of empirical research, readers should interpret these research areas as closely interconnected (e.g., developing PCK through teacher training also requires careful thought about the role of PCK in shaping teaching practice and student learning).

## Development of PCK

About half of the reviewed DRK-12 projects (14 of 27) investigated the development of teachers' PCK. Some projects examined the natural acquisition of PCK (e.g., in response to teaching practice; Carrier et al., 2018), but most were intervention studies that developed and tested PD programs aiming to improve teacher PCK. The intervention studies were generally small-scale, early-phase investigations focused on the initial design and development of new PD programs (e.g., Knudsen et al., 2015). This section reflects these development goals by featuring design considerations for fostering teacher PCK, though we also reviewed studies for evidence of causal impact (What Works Clearinghouse [WWC], 2020).

### Findings

As shown in Table 2, common PD design principles included (a) analyze student work, (b) embed active learning, and (c) situate in classroom contexts. Although not every intervention study included all three principles, researchers often considered them as key components to develop teachers' PCK, based on the theory of change and initial implementation evidence. For instance, Project Spotlight 2 details these considerations for the project, *Energy: A Multidisciplinary Approach for Teachers (EMAT): Designing and Studying a Multidisciplinary, Online Course for High School Teachers*.

**Table 2. Emerging Design Principles for Fostering Teacher PCK**

| Principle | Examples |
|---|---|
| Analyze student work | • Review videos of classroom discussions for evidence of student thinking (e.g., Kowalski et al., 2018; Williams & Clement, 2014; Wilson et al., 2017).<br>• Review written student work or written transcripts of simplified teacher-student interactions (e.g., Goldenberg et al., n.d.; Louie & Nikula, 2019; Mateas, 2016). |
| Embed active learning | • Ask pairs of teachers to defend to each other their analysis of student thinking and specific mathematics problems (e.g., Kowalski et al., 2018; White et al., 2013).<br>• Have the workshop facilitator engage teachers in group discussions about specific student thinking issues (e.g., Kowalski et al., 2018; Goldenberg et al., n.d.). |
| Situate in classroom contexts | • Use videos of actual classroom interactions to guide workshop discussions (e.g., Kowalski et al., 2018; Wilson et al., 2017).<br>• Use improvisational teaching games, plan through visualization exercises, and draft lesson plans (e.g., Knudsen et al., 2015; You & Kapila, 2017). |

PD programs that used these principles showed some initial promise in improving teachers' PCK (e.g., Knudsen et al., 2015; Kowalski et al., 2018). Project Spotlight 2 also presents some design challenges. Reflecting their design and development goals, these DRK-12 studies often did not aim to use rigorous causal designs. In several studies, teachers' PCK increased from before to after a PD intervention, but these pre-post designs provide weak causal evidence because change across time could result from other factors (e.g., retesting effects, natural growth of PCK; Marsden & Torgerson, 2012). Some studies used more rigorous experimental or quasi-experimental designs with comparison groups (e.g., Jacobs et al., 2019; Wilson et al., 2017). But even for those, the study reports generally lacked key information (e.g., attrition, baseline equivalence) needed to assess study quality. Our review of the study reports (see Table B2 in Appendix B) indicates that likely none would meet the U.S. Department of Education's standards for high-quality causal evidence (i.e., WWC Group Design Standards; WWC, 2020), except for one low-attrition randomized controlled trial (Jacobs et al., 2019).

## *Summary*

DRK-12 projects that investigated the development of teacher PCK were often early-phase intervention studies of PD programs. Results suggested several practical design considerations and challenges for fostering teacher PCK. This initial evidence provides opportunities for larger scale research evaluations, but DRK-12's current portfolio of causal evidence would not yet support the widespread adoption of the developed PD programs.

Teachers' Pedagogical Content Knowledge in Mathematics and Science

# Project Spotlight 2: Development of PCK

*Energy: A Multidisciplinary Approach for Teachers (EMAT): Designing and Studying a Multidisciplinary, Online Course for High School Teachers* (NSF award #1118643; funded amount = $3.1 million)

## Why Spotlight This Project?

This project used a mixed-methods approach, leveraging both quantitative and qualitative evidence, to understand how to adapt a previously successful PCK-related intervention to a new professional learning context: online PD spanning multiple science fields.

## What Was Studied?

The study developed and tested a 10-week, 120-hour online PD summer course that covered energy topics in high school science curricula. The course aimed to foster teachers' PCK by engaging them in analyzing video recordings of science teaching and learning. Two core theoretical principles drove the design of PD materials: active learning and situated cognition.

| Principle | Authors' description | PD implementation |
|---|---|---|
| Active learning | Active learning occurs when learners express their current understanding, examine new data, identify contradictions with expectations, and reflect on their revised understanding. | Teachers considered how they currently teach a topic, watched videos of student learning, discussed evidence of student ideas, and reflected on new ways to teach and understand student ideas. |
| Situated cognition | Knowledge does not exist independent of social contexts and its applications. Teachers should learn in the context of how they might teach. | Teachers watched videos of prior real-world science instruction about energy topics that teachers would likely teach in their own high school classrooms. |

The study also expanded the target audience of a previously successful PD intervention that used the same video-based analysis-of-practice model. These expansions included broadening the disciplinary focus (multiple science fields rather than a single field) and accessibility (online rather than face-to-face), but these changes also introduced considerable challenges.

## What Was Found?

Despite some positive teacher-level effects, both quantitative and qualitative evidence indicated challenges with expanding the intervention model. The authors suggested to improve the online-based implementation by providing more example lessons, extending scaffolding of the video analysis protocols, and spacing the PD during a full school year rather than one summer.

- **Quantitative evidence**. Although teachers' ability to analyze video clips of science learning (i.e., PCK-related measure) improved across time, the quantitative gains were about half the size as found in the earlier face-to-face implementation. Effects on student achievement were also much smaller.

- **Qualitative evidence**. Teachers' qualitative comments about the PD course revealed insights about the smaller quantitative gains. The researchers focused on five case study teachers. Although some teachers raved about the course, other teachers struggled to find the active learning opportunities helpful, desiring more feedback. Others failed to see connections between the PD course content and their classroom practice.

Teachers' Pedagogical Content Knowledge in Mathematics and Science

## Relationship of PCK to Teaching Practice

Improving teachers' PCK does not necessarily change teachers' behavior and classroom practice, though the constructs are closely related (e.g., Gess-Newsome, 2015). To empirically characterize this relationship, we identified DRK-12 projects that included distinct measures of PCK and teaching practice (e.g., delivered an MKT measure and conducted separate classroom observations). We found five such projects meeting this criterion, covering content areas such as mathematical argumentation in middle school (Knudsen et al., 2015), energy topics in science instruction (Kowalski et al., 2018), interactive simulations in algebra instruction (Findley et al., 2017), mathematics more broadly (Mosvold & Hoover, 2017), and science more broadly (Wilson et al., 2017).

This section does not cover PCK measurement approaches that observe PCK in action, such as in clinical simulations and lesson observations. Despite their affordances, these approaches make the distinction and relationship between PCK and practice more difficult to disentangle and study empirically. However, this section included studies that examined relationships between other PCK measures with teaching practice as measured by lesson observations.

### *Findings*

Project Spotlight 3, *Preparing Urban Middle Grades Mathematics Teachers to Teach Argumentation*, details a randomized experiment (*n* = 31 teachers) that examined the effects of a "Bridging PD" that explicitly connected MKT to future classroom practice through improvisational teaching games and planning through visualization (Knudsen et al., 2015). Based on the results detailed in Project Spotlight 2 about effects on teaching practice, the authors argued that intervention developers must be intentional in connecting lessons about PCK to concrete plans for future classroom practice, offering the "Bridging PD" as one example.

> Intervention developers must be intentional in connecting lessons about PCK to concrete plans for future classroom practice.

This point about building bridges to classroom practice aligns with the conclusions of a DRK-12-funded review of 12 empirical studies (NSF funded or not) about how MKT relates to mathematics teaching practice (Mosvold & Hoover, 2017). The review found that prior research has often focused on identifying the knowledge that teachers possess, but the authors argued that the field should focus more on detailing the work of doing mathematics teaching. Another mathematics DRK-12 project, studying interactive simulations in algebra instruction, found that individual differences in teachers' use of interactive simulations in the classroom (i.e., teacher practice) closely aligned with teachers' beliefs (i.e., PCK) about the affordances and drawbacks of using interactive simulations to teach mathematics concepts (Findley et al., 2017).

In the science domain, two DRK-12 projects examined teaching practice in the context of intervention research that trained in-service and preservice teachers to analyze video clips of science teaching and learning (Kowalski et al., 2018; Wilson et al., 2017). Although both projects included separate measures of teacher PCK and practice (see Table B3 in Appendix B), the study reports did not provide analyses explicitly relating these two constructs. The reported findings, however, provided some evidence that the PD intervention increased use of the taught instructional strategies.

## *Summary*

Although few DRK-12 projects included distinct measures of both PCK and teaching practice, projects in the mathematics domain indicated several notable findings. Study authors argued that MKT researchers should be intentional in connecting MKT to real-world teaching practice, both in terms of designing interventions and detailing the mathematical work of doing teaching.

Teachers' Pedagogical Content Knowledge in Mathematics and Science

# Project Spotlight 3: Relationship of PCK to Teaching Practice

*Preparing Urban Middle Grades Mathematics Teachers to Teach Argumentation* (NSF award #1417895; funded amount = $3 million)

## Why Spotlight This Project?

As a central focus, this study aimed to help teachers connect their MKT to concrete plans for teaching and supporting student argumentation.

## What Was Studied?

The evaluation design randomly assigned 31 middle school mathematics teachers to two alternate versions of a 2-week summer PD program.

| PD week | Treatment version | Control version |
|---|---|---|
| 1 | Sessions about MKT | Sessions about MKT |
| 2 | Bridging PD model | Alternate PD content |

In Week 1, both versions were identical and focused on MKT. But in Week 2, only the treatment version had a Bridging PD component that explicitly connected MKT to classroom practice through improvisational teaching games and planning through visualization.

## What Was Found?

- **Teaching Practices to Support Argumentation**. Treatment teachers more often engaged in teaching moves that were coded as supporting student argumentation, such as asking open-ended questions and encouraging the participation of multiple students.
  - Researchers studied these teaching practices by videotaping whole-classroom discussions and transcribing the teacher and student talk verbatim.
  - Observations were made across 2 days in which teachers taught using materials covered in Week 1 of the program.
- **Student Argumentative Talk**. The frequency of student argumentative talk was approximately twice as high in treatment than control classrooms across the two cohort years. Also, the average length of students' arguments correlated with teachers' MKT scores.

Although these findings more directly concern the empirical line of research about student outcomes, they also presumably reflect the impact of differing teaching practices.

### Conjecturing

- Look for math patterns that make sense to you
- Think about more than just one case
- Be creative
- Don't judge other people's conjectures

### Justifying

- Look for reasons why a conjecture is true of false
- Consider examples and counterexamples
- Build off of other people's ideas
- Generalize
- Try to convince others of your ideas, but keep in mind that you could be wrong…which is OK

### Concluding

- End an argument when your class is convinced that the conjecture is true (or false)

A teacher made a poster with these points during the PD workshop to make explicit the norms and behaviors expected from students when making mathematical arguments.

## Relationship of PCK to Student Learning

By developing their PCK, teachers can engage in classroom practices that better support learning, potentially improving student outcomes. We identified five DRK-12 projects with eligible award dates that examined the link between teacher PCK and student outcomes (see Table B4 in Appendix B). These outcomes included students' argumentative talk (Knudsen et al., 2015, Lee & Walkowiak, 2016), multiplicative reasoning (Tzur et al., 2018), energy content knowledge (Kowalski, 2016; Kowalski et al., 2018), and science achievement (Wilson et al., 2017). Study methods included both correlational analyses and student-level impact analyses of PCK-related PD programs.

### *Findings*

Some DRK-12 projects directly related individual differences in teachers' PCK to student outcomes. For instance, in one study, ratings of student explanation and justification discourse were higher for teachers with higher MKT scores (Lee & Walkowiak, 2016). The Project Spotlight 4, *Project ATOMS: Accomplished Elementary Teachers of Mathematics and Science*, details this study. These findings also conceptually align with those for another DRK-12 project on student argumentative talk in mathematics (Knudsen et al., 2015).

> Ratings of student explanation and justification discourse were higher for teachers with higher MKT scores.

However, not all projects found consistent evidence for such a teacher-student link. For instance, in one study ($n$ = 47 high school science teachers), teachers' abilities to analyze video clips of science learning did not significantly correlate with students' content knowledge about energy ($p$s > .63; Kowalski, 2016; Kowalski et al., 2018). Furthermore, teachers' PCK increased across time in that study, despite limited observable change in student science achievement. Also, the earlier-mentioned project on student mathematical discourse indicated mixed findings (i.e., significant correlations for some but not all ratings of student discourse; Lee & Walkowiak, 2016). These findings suggest nuance of when and how teachers' PCK shapes student learning, resonating with the earlier conclusion that improved PCK does not necessarily lead to improved teaching practice.

Lastly, several projects conducted simple between-group analyses to examine the student-level effects of PD programs that included PCK-related content (e.g., Kowalski et al., 2018; Wilson et al., 2017). These analyses, however, provide only indirect evidence of how teachers' PCK shapes student learning. For instance, one quasi-experimental study evaluated the impact of a PD program aiming to bolster third-grade teachers' understanding of students' multiplicative reasoning (Tzur et al., 2018). Although the study found evidence for positive effects on students' mathematics performance, the differences between intervention and comparison teachers could have reflected confounds besides teachers' PCK (e.g., learning broader teaching

Teachers' Pedagogical Content Knowledge in Mathematics and Science

practices distinct from topic-specific knowledge about student reasoning). A statistical tool known as *mediation analysis* (e.g., Lachowicz et al., 2018) could help address such limitations by examining how intervention effects on teachers' PCK explain student-level effects. However, the reviewed projects generally did not use this analytic approach.

## *Summary*

Few of the reviewed DRK-12 projects directly examined how teachers' PCK related to student learning. However, in two projects, student argumentative talk was rated higher in classrooms of teachers with higher MKT scores. Intervention studies indicated opportunities for using alternate analytic approaches (e.g., mediation analysis) that could combine analyses of intervention effects with correlational analyses of individual teacher differences.

# Project Spotlight 4: Relationship of PCK to Student Outcomes

*Project ATOMS: Accomplished Elementary Teachers of Mathematics and Science* (NSF award #1118894; total funded amount = $3.2 million)

## Why Spotlight This Project?

It used a relatively large sample of novice elementary school teachers (*n* = 118), administered previously validated measures of teacher mathematical knowledge and student classroom discourse, and conducted rigorous analyses using multilevel modeling with statistical controls for covariates.

## What Was Studied?

The project investigated how teachers' MKT in the number and operations domains related to classroom observations of student mathematical discourse.

## What Was Found?

- **Explanation and Justification**. Teachers' MKT scores positively correlated with ratings of student explanation and justification classroom discourse.
  - These ratings were based on classroom observations of the presence and depth of these constructs (see the below rubric). Although this rubric includes some teacher-based indicators (e.g., the types of questions teachers ask), most were student-based indicators.
  - The relationship between teachers' MKT and student discourse ratings remained even after statistically controlling for covariates such as teachers' self-efficacy, teachers' perceptions of school-level support for mathematics instruction, and school-level socioeconomic status.

- **Mathematical Discourse Community**. Teachers' MKT scores did not significantly correlate with ratings of mathematics discourse community.
  - These ratings were based on indicators for (a) teacher's role in discourse, (b) sense of mathematics community through student talk, and (c) questions.
  - In their interpretation, the authors suggested that "a teacher's level of MKT does not seem to influence their likelihood to solicit student ideas or allow opportunities for student-to-student talk, but it does impact the level of questioning posed to promote students' explanation of ideas" (Lee & Walkowiak, 2016, p. 1249).

## Example Classroom Observation Rubric (from Walkowiak et al., 2014)

6. Explanation and justification: The extent to which the teacher expects and students provide explanations/ justifications, both orally and on written assignments. NOTE: Simply "showing your work" on written assignments — i.e., writing the steps involved in calculating an answer — does not constitute an explanation.

| | Low (1, 2) | Medium (3, 4, 5) | High (6, 7) |
|---|---|---|---|
| Presence of explanation and justification | Students rarely provide explanations or justify their reasoning. | Students sometimes provide explanations and/or justify their reasoning. | Students often provide explanations and/or justify their reasoning. |
| | Teachers rarely ask "what, how, why" questions or otherwise solicit student explanations/justifications. | Teachers sometimes ask "what, how, why" questions or otherwise solicit student explanations/justifications. | Teachers often ask "what, how, why" questions or otherwise solicit student explanations/justifications. |
| Depth of explanation and justification (procedural and conceptual) | Student explanations often focus on procedural steps *and rarely* include conceptual understanding of the topic(s). | Student explanations sometimes focus on procedural steps *and sometimes* include conceptual understanding of the topic(s). | Student explanations rarely focus on procedural steps *and often* on conceptual understanding of the topic(s). |

# What Are the Key Takeaways?

This review synthesized insights from 27 NSF-funded projects, totaling $62 million awarded, that studied PCK in PK–12 STEM education, roughly equally split across mathematics and science education. The projects primarily applied correlational/observational and longitudinal methods, often targeted teaching in the middle school grades, and used a variety of approaches to measure teachers' PCK. The projects advanced substantive knowledge about PCK across four major lines of research, especially regarding the measurement and development of PCK.

## Measurement of PCK

One major cross-cutting contribution was methodological advances in the approaches to measure PCK. In mathematics education, some DRK-12 projects innovated on the classic MKT assessments, such as developing new measures for using visual representations (DePiper & Driscoll, 2018; Louie & Nikula, 2019) or using qualitative interviews to better understand existing MKT items (Hoover & Lai, 2017).

> The DRK-12 projects expanded the toolkit for investigating PCK in PK–12 STEM education through advances in measurement.

The history of using validated PCK measures has been shorter in science education research than mathematics education research. But the DRK-12 projects made key methodological advances in the science domain too, such as building measures to assess teachers' content knowledge for teaching energy (Etkina et al., 2018) or PCK about scientific argumentation (Katsh-Singer et al., 2016; McNeill et al., 2014).

Both mathematics- and science-focused DRK-12 projects identified key affordances and challenges of various measurement approaches. For instance, projects highlighted how vignettes can ground multiple-choice PCK assessment items in specific pedagogical contexts, but picking these vignettes requires careful design and empirical testing (Hoover & Lai, 2017; McNeill et al., 2014; Wittmann et al., 2017).

Other projects innovated on qualitative approaches to capture teachers' PCK, such as through qualitative interviews (e.g., Smith et al., 2017) or observing PCK in action through lesson observations (e.g., Robertson et al., 2017). One emerging approach was clinical simulations that involve interactions with "standardized" students; this approach aims to observe PCK in action but in more controlled settings than classroom teaching (Dotger et al., 2015; Dotger et al., 2018; Shaughnessy & Boerst, 2018; Shaughnessy et al., 2018).

Overall, the DRK-12 projects expanded the methodological toolkit for investigating PCK in PK–12 STEM education, especially in mathematics and science education. Selecting a specific

Teachers' Pedagogical Content Knowledge in Mathematics and Science

measurement approach requires aligning the research goal with the method's affordances. As several study authors suggested, including multiple measurement approaches can help combine the strengths of each (e.g., using cognitive interviews to understand how teachers respond to multiple-choice items; Hoover & Lai, 2017). As illustrated in mathematics education research, such advances in measurement are key to fostering the capacity to conduct research on PCK and generate other substantive insights (Depeape et al., 2013).

## Development of PCK

Another major strand of research in these DRK-12 projects was the development of PCK, especially through PD programs for in-service teachers. The intervention studies were generally early-phase investigations, often focused on designing and refining new PD programs. The studies indicated some common design principles for fostering teachers' PCK, such as (a) analyze student work, (b) engage teachers in active learning, and (c) situate the PD in classroom contexts.

PD programs that employed these design principles showed some promise of improving teachers' PCK. However, limitations in the evaluation designs and study reporting prevented our synthesis from drawing any strong causal conclusions. For instance, among the reviewed DRK-12 projects, we did not find any evaluation of PCK-related interventions that would have met WWC standards for high-quality causal evidence based on the study report alone, except for one low-attrition randomized controlled trial (Jacobs et al., 2019).

The limitations in causal evidence likely reflect these projects' design and development goals (e.g., creating prototypes and proofs of concept). In other words, conducting large-scale randomized controlled trials was premature for the exploratory goals of many of these projects. Nevertheless, these considerations suggest opportunities for causal evaluation in future research.

## Relationships With Teaching Practice and Student Learning

Compared with projects characterizing and improving teachers' PCK, fewer investigated the relationship between teachers' PCK with teaching practice and student learning. However, the projects that investigated these relationships indicated the value of doing so.

One key message was that PD invention developers should be intentional in connecting lessons about PCK to concrete plans for improving classroom practice and student learning (Mosvold & Hoover, 2017). For instance, the *Preparing Urban Middle Grades Mathematics Teachers to Teach Argumentation* project (see Project Spotlight 3) investigated the concept of a "bridging" model that explicitly connected PD lessons about PCK to classroom practice through improvisational teaching games and planning through visualization (Knudsen et al., 2015). The

Teachers' Pedagogical Content Knowledge in Mathematics and Science

project's randomized experiment found that this bridging component was essential to foster teachers' use of practices to support student argumentation in the classroom.

Projects that investigated these relationships tended to do so in two main ways: (a) correlations between PCK with teaching practice and student outcomes and (b) effects of PCK-related intervention on student learning outcomes. Results from the second approach were ambiguous because the study design did not isolate the unique role of teacher's PCK specifically. These considerations suggest opportunities for future research to better understand the mechanisms of how PCK shapes teaching practice and student learning (and how teaching experiences reciprocally enhance PCK).

## Opportunities for Future Research

In this section, we consider opportunities and priorities for future research, critically examining potential gaps in the DRK-12 portfolio. We focus on ways that the construct of PCK can help future NSF-funded researchers advance DRK-12's central mission to enhance STEM teaching and learning in grades PK–12. Our commentary centers on intervention research (e.g., developing and evaluating PD programs) or research that directly informs intervention studies (e.g., creating PCK measures that can serve as outcomes). We organized our reflections[3] based on three main categories: (a) aligning theoretical and measurement models of PCK, (b) developing causal evidence, and (c) understanding mechanisms of change. These three categories map onto the major lines of PCK research reviewed earlier (the nature of PCK, the development of PCK, and the relationship to teaching practice and student learning).

### *Aligning Theoretical and Measurement Models*

The methodological advances in PCK measurement offer several opportunities for using the construct of PCK to improve STEM teaching (Uzzo et al., 2018). For instance, assessments of PCK can allow teachers to understand PCK concretely and diagnose limitations in their current professional knowledge base. Moreover, intervention studies require validated PCK assessments if they aim to empirically understand how to improve teachers' PCK.

Taking advantage of these opportunities, however, also requires aligning the measurement models with the theoretical models of PCK. Our report started by reviewing some of the major controversies in defining and conceptualizing PCK in mathematics education (e.g., Depaepe et al., 2013) and science education (e.g., Gess-Newsome, 2015). Part of the approach to resolve these controversies has been to recognize the multifaceted nature of PCK. For instance, the refined consensus model in science education (Carlson et al., 2019) describes three kinds of PCK

---

[3] Two external content experts also contributed significantly to these synthesis reflections: Stephen Uzzo at the New York Hall of Science (expert on PCK in science) and Paola Sztajn at North Carolina State University (expert on PCK in mathematics). However, this section does not necessarily reflect their endorsement of our reflections.

Teachers' Pedagogical Content Knowledge in Mathematics and Science

(enacted, personal, and collective) while also making further distinctions with those broad categories (e.g., grain size at the discipline, topic, and concept levels). Other distinctions between types of PCK include components such as knowledge about student thinking versus instructional strategies versus curriculum (Depaepe et al., 2013). Furthermore, complexities in content knowledge for teaching frameworks (e.g., MKT) include distinguishing between PCK and specialized content knowledge (Petrou & Goulding, 2011).

The multifaceted nature of PCK offers another way to contrast measurement approaches. For instance, lesson observations and clinical simulations would be well suited for studying enacted PCK (i.e., as used in specific teaching scenarios), whereas quantitative tests of PCK are suited for studying personal PCK (i.e., a teacher's reservoir of knowledge). Hence, one goal for future research would be to draw from emerging conceptual models of PCK (e.g., Carlson et al., 2019) and systematically map their components to concrete PCK measures and measurement best practices (e.g., Liepertz & Borowski, 2019).

Given practical constraints, any individual project likely will not be able to attend to and measure all possible aspects of PCK, even if the project constrains the content area focus (e.g., limiting to PCK about photosynthesis). Such limitations are sensible from the perspective of an individual researcher. Nevertheless, these constraints pose considerable challenges for the field when developing unifying principles to link the small facets of PCK examined in individual studies to a united whole. Progress in the field will likely come from both top-down approaches (e.g., conceptual models guiding new research) and bottom-up approaches (e.g., such as this synthesis of past empirical research).

### *Developing Causal Evidence*

The reviewed DRK-12 portfolio produced promising proofs of concept but limited causal evidence on what interventions can improve teachers' PCK. For any specific funded project, such limitations may appropriately reflect the scope and goals of early-phase intervention design research. However, we see cause for concern if the aggregate portfolio of funded work has major limitations in providing larger scale, causal evidence. Such limitations suggest that this portfolio of DRK-12 research may be many years away from using the concept of PCK to advance DRK-12's mission to significantly enhance PK–12 STEM teaching and learning at a national scale.

Using research to enhance educational practice likely requires an ecosystem of coexisting, distinct research types that mutually inform each other, as described in NSF's and the Institute for Education Sciences' (2013) *Common Guidelines for Education Research and Development*. In this respect, we do not wish to criticize specific exploratory or development projects, for which

Teachers' Pedagogical Content Knowledge in Mathematics and Science

providing strong evidence of effectiveness is not a reasonable expectation. Rather, we note that such initial design research creates the preconditions that warrant studies on intervention efficacy and scale-up (Sztajn et al., 2017). The time may be especially ripe for more DRK-12 impact studies on promising interventions for enhancing teachers' PCK.

In some cases, improving the causal evidence base could be addressed through better reporting practices, such as providing key information on attrition in randomized designs and baseline equivalence in quasi-experimental designs. In other cases, the study designs themselves would need improvement. For instance, many reviewed studies simply examined change across time in one group of teachers, with no comparison group of teachers to help rule out alternative explanations. Only three of the 27 projects we reviewed included a randomized experiment.

## Understanding Mechanisms of Change

Although many projects studied PCK-related interventions, far fewer examined how improving PCK can affect teaching practice and student learning. This limitation is especially important because the projects with relevant evidence showed that the relationships between PCK, teaching practice, and student learning are far from straightforward.

In some cases, the available data in these projects indicated opportunities for more sophisticated modeling approaches that could address some of these limitations. Some projects evaluated the effects of PD programs separately for PCK, teaching practice, and/or student learning outcomes. Although separately examining these outcomes is a useful starting point, that approach does not directly shed light on the unique role of PCK in the mechanisms of change. For instance, if an intervention included both PD about PCK and new curriculum materials, any changes in teaching practice or student learning could result from the new curriculum materials rather than improvements in PCK.

A statistical tool known as mediation analysis (MacKinnon et al., 2007) could address such limitations by investigating how improvements in PCK mediate changes in teaching practice and student learning. Notably, some reviewed projects already have the data to apply this approach (e.g., intervention studies that already measured both PCK and teaching practice). Mediation analysis is not a panacea, but it can offer insight on mechanisms of change. Mediation models also offer ways to move beyond simple bivariate correlations in observational, nonintervention research, allowing the researcher to specify structural relationships between constellations of constructs and compare model fit with alternate model specifications (Preacher, 2006).

Investigating PCK within a broader constellation of relevant teacher and student constructs can offer several benefits for other lines of PCK research as well. For instance, consider if an intervention study finds that measures of PCK did not mediate changes in teaching practice and student learning. That finding could both prompt changes in the measurement of PCK (e.g., perhaps the measure was not well aligned with the types of PCK needed in real-world classroom instruction) and in interventions to develop PCK (e.g., perhaps the lessons about PCK were not well tailored to real-world instruction).

## Conclusions

Teacher PCK is a complex, multifaceted construct that is widely seen as foundational to the act of teaching. Our synthesis focused on how recent DRK-12 projects have studied PCK, including its measurement, development, and relationship to teaching and student learning. The recent advancements in measurement approaches make PCK a topic ripe for further exploration, especially in PK–12 mathematics and science education. Several intervention studies yielded promising proofs of concept and design principles for improving teachers' PCK, providing opportunities for larger scale causal evaluation in future research. Limitations in the reviewed DRK-12 portfolio also suggest opportunities for studying how PCK fits into a complex system of broader teacher professional knowledge, teacher practice, and student learning. Future intervention research, for instance, could advance understanding about mechanisms of change by studying PCK as a mediator of intervention effects on teaching practice and student learning. Pursuing such methodological, theoretical, and empirical advances about PCK could help support DRK-12's mission to significantly enhance PK–12 STEM teaching and learning.

# References

Abell, S. K. (2008). Twenty years later: Does pedagogical content knowledge remain a useful idea? *International Journal of Science Education*, *30*(10), 1405–1416. https://doi.org/10.1080/09500690802187041

Ball, D. L., & Cohen, D. K. (1999). Developing practice, developing practitioners: Toward a practice-based theory of professional education. In G. Sykes & L. Darling-Hammond (Eds.), *Teaching as the learning profession: Handbook of policy and practice* (pp. 3–32). Jossey-Bass.

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, *59*(5), 389–407. https://doi.org/10.1177/0022487108324554

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Klusmann, U., Neubrand, M., & Tsai, Y.-M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, *47*(1), 133–180. https://doi.org/10.3102/0002831209345157

Bragelman, J., Stoelinga, T., & Castro Superfine, A. (2017). Iterative (re)visioning: An improvement science approach to online professional development design and implementation. In E. Galindo & J. Newton (Eds.), *Proceedings of the 39th annual meeting of the North American Chapter of the Psychology of Mathematics Education* (p. 1379). Hoosier Association of Mathematics Teacher Educators.

Brill, A., Elliott, C. H., Listman, J. B., Milne, C. E., & Kapila, V. (2016). Middle school teachers' evolution of TPACK understanding through professional development. In *Proceedings of the ASEE Annual Conference & Exposition* (Vol. 18260, p. 25720). https://peer.asee.org/25720.pdf

Carlson, J., Daehler, K. R., Alonzo, A. C., Barendsen, E., Berry, A., Borowski, A., Carpendale, J., Kam Ho Chan, K., Cooper, R., Friedrichsen, P., Gess-Newsome, J., Henze-Rietveld, I., Hume, A., Kirschner, S., Liepertz, S., Loughran, J., Mavhunga, E., Neumann, K., Nilsson, P., Park, S., . . ., Wilson, C. D. (2019). The refined consensus model of pedagogical content knowledge in science education. In A. Hume, R. Cooper, & A. Borowski (Eds.), *Repositioning pedagogical content knowledge in teachers' knowledge for teaching science* (pp. 77–94). Springer. https://doi.org/10.1007/978-981-13-5898-2_2

Carnegie–Institute for Advanced Study Commission on Mathematics and Science Education. (2009). *The opportunity equation: Transforming mathematics and science education for citizenship and the global economy*. https://www.carnegie.org/media/filer_public/80/c8/80c8a7bc-c7ab-4f49-847d-1e2966f4dd97/ccny_report_2009_opportunityequation.pdf

Carrier, S. J., Whitehead, A. N., Minogue, J., & Corsi-Kimble, B. S. (2018). Novice elementary teachers' developing visions of effective science teaching. *Research in Science Education*, *50*(4), 1521–1545. https://doi.org/10.1007/s11165-018-9742-7

Castro Superfine, A., & Li, W. (2017). Characterizing teachers' informal conceptions of learning trajectories in mathematics. In E. Galindo & J. Newton (Eds.), *Proceedings of the 39th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education: Synergy at the Crossroads: Future Directions for Theory, Research, and Practice* (p. 1259). Hoosier Association of Mathematics Teacher Educators. https://www.pmena.org/pmenaproceedings/PMENA%2039%202017%20Proceedings.pdf

Chan, K. K. H., & Hume, A. (2019). Towards a consensus model: Literature review of how science teachers' pedagogical content knowledge is investigated in empirical studies. In A. Hume, R. Cooper, & A. Borowski (Eds.), *Repositioning pedagogical content knowledge in teachers' knowledge for teaching science* (pp. 3–76). Springer. https://doi.org/10.1007/978-981-13-5898-2_1

Copur-Gencturk, Y., & Lubienski, S. T. (2013). Measuring mathematical knowledge for teaching: A longitudinal study using two measures. *Journal of Mathematics Teacher Education*, *16*(3), 211–236. https://doi.org/10.1007/s10857-012-9233-0

Crismond, D., & Adams, R. (2012). The informed design teaching and learning matrix. *Journal of Engineering Education*, *101*(4), 738–797. https://doi.org/10.1002/j.2168-9830.2012.tb01127.x

Crismond, D., & Lomask, M. (2016, June 26–29). *Using engineering design teaching portfolios to gauge design teacher performance and infer design pedagogical content knowledge* [Paper presentation]. 2016 ASEE Annual Conference & Exposition, New Orleans, LA, United States. https://peer.asee.org/using-engineering-design-teaching-portfolios-to-gauge-design-teacher-performance-and-infer-design-pedagogical-content-knowledge

Darling-Hammond, L. (n.d.). *Improving competency in elementary science teaching and learning (NSF project)*. Stanford Center for Assessment, Learning and Equity. https://scale.stanford.edu/content/improving-competency-elementary-science-teaching-and-learning-nsf-project

Depaepe, F., Verschaffel, L., & Kelchtermans, G. (2013). Pedagogical content knowledge: A systematic review of the way in which the concept has pervaded mathematics educational research. *Teaching and Teacher Education*, *34*, 12–25. https://doi.org/10.1016/j.tate.2013.03.001

DePiper, J. N., & Driscoll, M. (2018). Teacher knowledge and visual access to mathematics. In S. M. Uzzo, S. B. Graves, E. Shay, M. Harford, & R. Thompson (Eds.), *Pedagogical content knowledge in STEM* (pp. 83–102). Springer. https://www.springer.com/gp/book/9783319974743

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, *38*(3), 181–199. https://doi.org/10.3102/0013189X08331140

Dotger, B., Masingila, J., Bearkland, M., & Dotger, S. (2015). Exploring iconic interpretation and mathematics teacher development through clinical simulations. *Journal of Mathematics Teacher Education*, *18*(6), 577–601. https://doi.org/10.1007/s10857-014-9290-7

Dotger, B., Dotger, S., Masingila, J., Rozelle, J., Bearkland, M., & Binnert, A. (2018). The right "fit": Exploring science teacher candidates' approaches to natural selection within a clinical simulation. *Research in Science Education*, *48*(3), 637–661. https://doi.org/10.1007/s11165-016-9582-2

Etkina, E., Gitomer, D., Iaconangelo, C., Phelps, G., Seeley, L., & Vokos, S. (2018). Design of an assessment to probe teachers' content knowledge for teaching: An example from energy in high school physics. *Physical Review Physics Education Research*, *14*(1), Article 010127-1. https://doi.org/10.1103/PhysRevPhysEducRes.14.010127

Evens, M., Elen, J., & Depaepe, F. (2015). Developing pedagogical content knowledge: Lessons learned from intervention studies. *Education Research International*, 2015, Article 790417. https://doi.org/10.1155/2015/790417

Fernandez, C. (2014). Knowledge base for teaching and pedagogical content knowledge (PCK): Some useful models and implications for teachers' training. *Problems of Education in the 21st Century*, *60*, 79–100. http://oaji.net/articles/2015/457-1421876658.pdf

Findley, K., Whitacre, I., & Hensberry, K. (2017). Integrating interactive simulations into the mathematics classroom: Supplementing, enhancing, or driving? [Paper presentation]. *Proceedings of the 39th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education: Synergy at the Crossroads: Future Directions for Theory, Research, and Practice* (pp. 1297–1304). Hoosier Association of Mathematics Teacher Educators. https://www.pmena.org/pmenaproceedings/PMENA%2039%202017%20Proceedings.pdf

Gess-Newsome, J. (2015). A model of teacher professional knowledge and skill including PCK: Results of the thinking from the PCK Summit. In A. Berry, P. Friedrichsen, & J. Loughran (Eds.), *Re-examining pedagogical content knowledge in science education* (pp. 28–42). Routledge.

Griffin, C. C., Jitendra, A. K., & League, M. B. (2009). Novice special educators' instructional practices, communication patterns, and content knowledge for teaching mathematics. *Teacher Education and Special Education: The Journal of the Teacher Education Division of the Council for Exceptional Children*, *32*(4), 319–336. https://doi.org/10.1177/0888406409343540

Haddaway, N. R., Collins, A. M., Coughlin, D., & Kirk, S. (2015). The role of Google Scholar in evidence reviews and its applicability to grey literature searching. *PloS One*, *10*(9), Article e0138237. https://doi.org/10.1371/journal.pone.0138237

Heller, J. I., Daehler, K. R., Wong, N., Shinohara, M., & Miratrix, L. W. (2012). Differential effects of three professional development models on teacher knowledge and student achievement in elementary science. *Journal of Research in Science Teaching*, *49*(3), 333–362. https://doi.org/10.1002/tea.21004

Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Education Research Journal*, *42*(2), 371–406. https://doi.org/10.3102/00028312042002371

Hill, H. C., Schilling, S. G., & Ball, D. L. (2004). Developing measures of teachers' mathematics knowledge for teaching. *The Elementary School Journal*, *105*(1), 11–30. https://www.jstor.org/stable/10.1086/428763

Hodkowski, N. M. (2018). *Manifestations of elementary mathematics teachers' shift towards second-order models* (Publication 10842329) [Doctoral dissertation, University of Colorado at Denver]. ProQuest Dissertations and Theses Global.

Hoover, M., & Lai, Y. (2017). Pedagogical context and its role in articulating mathematical knowledge for teaching. In S. Zehetmeier, B. Rösken-Winter, D. Potari, & M. Ribeiro (Eds.), 2017 *Proceedings of the Third ERME Topic Conference on Mathematics Teaching, Resources and Teacher Professional Development* (pp. 167–176). https://hal.archives-ouvertes.fr/ETC3/public/Full_Download.pdf

Hoover, M., Mosvold, R., Ball, D. L., & Lai, Y. (2016). Making progress on mathematical knowledge for teaching. *The Mathematics Enthusiast*, *13*(1–2), 3–34. https://scholarworks.umt.edu/tme/vol13/iss1/

Institute of Education Sciences and National Science Foundation. (2013). *Common guidelines for education research and development*. https://www.nsf.gov/pubs/2013/nsf13126/nsf13126.pdf

Jacobs, J. K., Koellner, K., Seago, N., Garnier, H., & Wang, C. (2019, April). *Context-driven learning from professional development: Examining teacher knowledge, instructional practice, and student mathematics achievement gains* [Paper presentation]. Annual conference of the American Educational Research Association, Toronto, Canada. https://www.aera.net/Publications/Online-Paper-Repository/AERA-Online-Paper-Repository

Jacobs, V. R., Empson, S. B., Krause, G. H., & Pynes, D. (2015). *Responsive teaching with fractions*. 2015 National Council of Teachers of Mathematics (NCTM) Research Conference, Boston, MA, United States. https://nctm.confex.com/nctm/2015RP/webprogram/Manuscript/Session33523/jacobs_empson_krause_pynes_NCTM2015_final.pdf

Kaarstein, H. (2014). A comparison of three frameworks for measuring knowledge for teaching mathematics. *Nordic Studies in Mathematics Education*, *19*(1), 23–52. http://ncm.gu.se/wp-content/uploads/2020/06/19_1_023052_kaarstein.pdf

Katsh-Singer, R., McNeill, K. L., & Loper, S. (2016). Scientific argumentation for all? Comparing teacher beliefs about argumentation in high, mid, and low socioeconomic status schools. *Science Education*, *100*(3), 410–436. https://doi.org/10.1002/sce.21214

Knudsen, J., Shechtman, N., Lara-Meloy, T., Stevens, H., & Rutstein, D. (2015). *A professional development program for mathematical argumentation: Bridging from workshop to classroom*. SRI International. https://bridgingmath.org/wp-content/uploads/2020/01/bridging-pd-for-math-argumentation.pdf

Kowalski, S. M. (2016). *Energy: A multidisciplinary approach for teachers (EMAT), research results* [Unpublished manuscript]. Education Development Center.

Kowalski, S. M., Stennett, B., Bloom, M., & Askinas, K. (2018). *Investigation of video-based multidisciplinary online professional development for inservice high school science teachers* [Unpublished manuscript]. Education Development Center.

Krause, G., Empson, S., Pynes, D., & Jacobs, V. (2016). Teachers' knowledge of children's strategies for equal sharing fraction problems. In *Proceedings of the 38th North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 466–473). The University of Arizona. http://www.pmena.org/pmenaproceedings/PMENA%2038%202016%20Proceedings.pdf

Lachowicz, M. J., Preacher, K. J., & Kelley, K. (2018). A novel measure of effect size for mediation analysis. *Psychological Methods*, *23*(2), 244–261. https://doi.org/10.1037/met0000165

Lee, I. A. (2018). *Developing teachers as computational thinkers through supported authentic experiences in computing modeling and simulation* [Annual project report]. National Science Foundation.

Lee, C. W., & Walkowiak, T. A. (2016, November 4–5). *Teacher attributes and school context: What are the barriers to discourse in elementary mathematics?* [Paper presentation]. 38th Annual Conference of the North American Chapter of the International Group for the Psychology of Mathematics Education, Tucson, AZ, United States. https://eric.ed.gov/?id=ED583793

Liepertz, S., & Borowski, A. (2018). Testing the consensus model: Relationships among physics teachers' professional knowledge, interconnectedness of content structure and student achievement. *International Journal of Science Education*, *41*(7), 890–910. https://doi.org/10.1080/09500693.2018.1478165

Louie, J., & Nikula, J. (2019). *Impact of professional development for math teachers of English learners* [Oral presentation]. Annual meeting for the National Council on Teaching Mathematics Research Conference, San Diego, CA, United States.

Lyon, E. G., Stoddart, T., Solis, J., Tolbert, S., Bunch, G. C., Roth, B., Salinas, I., Knox, C., Couling, J., & Butler, M. (2016). *Teaching English learners through science-language integration: Linking a conceptual framework to secondary teacher preparation* [Symposium]. Annual meeting of the National Association for Research on Science Teaching, Baltimore, MD, United States. https://www.researchgate.net/publication/301494737_Teaching_English_Learners_through_Science-Language_Integration_Linking_a_Conceptual_Framework_to_Secondary_Teacher_Preparation_Symposium_presented_at_the_annual_meeting_or_the_National_Association_for/link/589cc40f45851573881bf667/download

MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annual Review of Psychology*, *58*(1), 593–614. https://doi.org/10.1146/annurev.psych.58.110405.085542

Marsden, E., & Torgerson, C. J. (2012). Single group, pre- and post-test research designs: Some methodological concerns. *Oxford Review of Education*, *38*(5), 583–616. https://doi.org/10.1080/03054985.2012.731208

Mateas, V. (2016). Debunking myths about the standards for mathematical practice. *Mathematics Teaching in the Middle School*, *22*(2), 92–99. https://doi.org/10.5951/mathteacmiddscho.22.2.0092

McNeill, K. L., Gonzalez-Howard, M., Katsh-Singer, R., & Loper, S. (2014, March). *Measuring pedagogical content knowledge of argumentation through the development of a teacher argumentation assessment* [Paper presentation]. Annual Meeting of the National Association for Research in Science Teaching, Pittsburgh, PA, United States. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.716.978&rep=rep1&type=pdf

Mosvold, R., & Hoover, M. (2017). Mathematical knowledge for teaching and the teaching of mathematics. In T. Dooley & G. Gueudet (Eds.), *Proceedings of the 10th Congress of European Research in Mathematics Education* (pp. 3105–3112). Dublin City University, Institute of Education. https://hal.archives-ouvertes.fr/hal-01949100

Parks, A. (n.d.). *Smarter together working conference: Developing a shared curriculum of complex instruction for elementary mathematics methods courses*. Community for Advancing Discovery Research in Education. https://cadrek12.org/projects/smarter-together-working-conference-developing-shared-curriculum-complex-instruction-elemen

Petrou, M., & Goulding, M. (2010). Conceptualising teachers' mathematical knowledge in teaching. In T. Rowland & K. Ruthven (Eds.), *Mathematical knowledge in teaching* (pp. 9–25). Springer. https://doi.org/10.1007/978-90-481-9766-8_2

Preacher, K. J. (2006). Quantifying parsimony in structural equation modeling. *Multivariate Behavioral Research*, *41*(3), 227–259. https://doi.org/10.1207/s15327906mbr4103_1

Rahman, S. M., Krishnan, J. V., & Kapila, V. (2017, June 25–28). *Exploring the dynamic nature of TPACK framework in teaching STEM using robotics in middle school classrooms* [Paper presentation]. 2017 ASEE Annual Conference and Exposition, Columbus, OH, United States. https://peer.asee.org/exploring-the-dynamic-nature-of-tpack-framework-in-teaching-stem-using-robotics-in-middle-school-classrooms.pdf

Robertson, A. D., Scherr, R. E., Goodhew, L. M., Daane, A. R., Gray, K. E., & Aker, L. B. (2017). Identifying content knowledge for teaching energy: Examples from high school physics. *Physical Review Physics Education Research*, 13, Article 010105. https://doi.org/10.1103/PhysRevPhysEducRes.13.010105

Roth, K. J., Garnier, H. E., Chen, C., Lemmens, M., Schwille, K., & Wickler, N. I. (2011). Videobased lesson analysis: Effective science PD for teacher and student learning. *Journal of Research in Science Teaching*, *48*(2), 117–148. https://doi.org/10.1002/tea.20408

Roth, K., Stuhlsatz, M., Kruse, R., & Wilson, C. (2018). *Videocases for science teaching analysis plus (ViSTA Plus): Efficacy of a videocase-based, analysis-of-practice teacher preparation program*. https://www.researchgate.net/grant/Videocases-for-Science-Teaching-Analysis-Plus-ViSTA-Plus-Efficacy-of-a-Videocase-based-Analysis-of-Practice-Teacher-Preparation-Program_3054406

Rowland, T., & Ruthven, K. (2011). Introduction: Mathematical knowledge in teaching. In *Mathematical knowledge in teaching* (pp. 1–5). Springer. https://doi.org/10.1007/978-90-481-9766-8_1

Schneider, R. M., & Plasman, K. (2011). Science teacher learning progressions: A review of science teachers' pedagogical content knowledge development. *Review of Educational Research*, *81*(4), 530–565. https://doi.org/10.3102/0034654311423382

Seago, N., Koellner, K., & Jacobs, J. (2018). Video in the middle: Purposeful design of video-based mathematics professional development. *Contemporary Issues in Technology & Teacher Education*, *18*(1), 29–49. https://www.learntechlib.org/primary/p/174330/

Seymour, J. R., & Lehrer, R. (2006). Tracing the evolution of pedagogical content knowledge as the development of interanimated discourses. *The Journal of the Learning Sciences*, *15*(4), 549–582. https://doi.org/10.1207/s15327809jls1504_5

Shaughnessy, M., & Boerst, T. A. (2018). Uncovering the skills that preservice teachers bring to teacher education: The practice of eliciting a student's thinking. *Journal of Teacher Education*, *69*(1), 40–55. https://doi.org/10.1177/0022487117702574

Shaughnessy, M., Boerst, T. A., & Farmer, S. O. (2018). Complementary assessments of prospective teachers' skill with eliciting student thinking. *Journal of Mathematics Teacher Education*, *22*(6), 607–638. https://doi.org/10.1007/s10857-018-9402-x

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, *15*(2), 4–14. https://doi.org/10.3102/0013189x015002004

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, *57*(1), 1–22. https://doi.org/10.17763/haer.57.1.j463w79r56455411

Singer, J. E., Ross, J. M., & Jackson-Lee, Y. (2016). Professional development for the integration of engineering in high school STEM classrooms. *Journal of Pre-College Engineering Education Research (J-PEER)*, *6*(1). https://docs.lib.purdue.edu/jpeer/vol6/iss1/3/

Smith, P. S., Esch, R. K., Hayes, M. L., & Plumley, C. L. (2016). *Developing and testing a method for collecting and synthesizing pedagogical content knowledge*. Horizon Research, Inc. http://www.horizon-research.com/horizonresearchwp/wp-content/uploads/2016/04/ASSET-NARST-paper.pdf

Smith, P. S., Plumley, C. L., & Hayes, M. L. (2017). *Eliciting elementary teachers' PCK for the small particle model*. Horizon Research. https://eric.ed.gov/?id=ED587205

Smith, P. S., Plumley, C. L., Hayes, M. L., & Esch, R. K. (2018). Personal and canonical PCK: A synergistic relationship? In S. M. Uzzo, S. B. Graves, E. Shay, M. Harford, & R. Thompson (Eds.), *Pedagogical content knowledge in STEM* (pp. 49–66). Springer. https://link.springer.com/chapter/10.1007/978-3-319-97475-0_3

Spencer, D. (n.d.). *Implementing the Mathematical Practice Standards: Enhancing teachers' ability to support the Common Core State Standards* [Video]. Education Development Center. https://vimeo.com/220748774

Sztajn, P., Borko, H., & Smith, T. (2017). Research on mathematics professional development. In J. Cai (Ed.), *Compendium for research in mathematics education* (pp. 793–823). National Council of Teachers of Mathematics.

Taylor, J., Roth, K., Wilson, C., Stuhlsatz, M., & Tipton, E. (2017). The effect of an analysis-of-practice, videocase-based, teacher professional development program on elementary students' science achievement. *Journal for Research on Educational Effectiveness*, *10*(2), 241–271. https://doi.org/10.1080/19345747.2016.1147628

Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology*, *8*(1), Article 45. https://doi.org/10.1186/1471-2288-8-45

Thomson, M. M., DiFrancesca, D., Carrier, S., & Lee, C. (2016). Teaching efficacy: Exploring relationships between mathematics and science self-efficacy beliefs, PCK and domain knowledge among preservice teachers from the United States. *Teacher Development*, *21*(1), 1–20. https://doi.org/10.1080/13664530.2016.1204355

Tirosh, D., Tsamir, P., Levenson, E., & Tabach, M. (2011). From preschool teachers' professional development to children's knowledge: Comparing sets. *Journal of Mathematics Teacher Education*, *14*(2), 113–131. https://doi.org/10.1007/s10857-011-9172-1

Tzur, R., Johnson, H. L., Hodkowski, N. M., Jorgensen, C., Nathenson-Mejia, S., Wei, B., Smith, A., & Davis, A. (2018). Impact of a student-adaptive pedagogy PD program on students' multiplicative reasoning. In T. E. Hodges, G. J. Roy, & A. M. Tyminski (Eds.), *Proceedings of the 40th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 1084–1090). http://www.pmena.org/pmenaproceedings/PMENA%2040%202018%20Proceedings.pdf

Uzzo, S. M., Graves, S. B., Shay, E., Harford, M., & Thompson, R. (Eds.). (2018). *Pedagogical content knowledge in STEM: Research to practice*. Springer.

Walkowiak, T. A., Berry, R. Q., Meyer, J. P., Rimm-Kaufman, S. E., & Ottmar, E. R. (2014). Introducing an observational measure of standards-based mathematics teaching practices: Evidence of validity and score reliability. *Educational Studies in Mathematics*, *85*(1), 109–128. https://doi.org/10.1007/s10649-013-9499-x

Wayne, A. J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research*, *73*(1), 89–122. https://doi.org/10.3102/00346543073001089

Westgate, M. J. (2018). revtools: An R package to support article screening for evidence synthesis. *Computational Tools and Methods*, *10*(4), 606–614. https://doi.org/10.1002/jrsm.1374

What Works Clearinghouse. (2020). *What Works Clearinghouse standards handbook, version 4.1*. U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. https://ies.ed.gov/ncee/wwc/handbooks

White, D., Donaldson, B., Hodge, A., & Ruff, A. (2013). Examining the effects of math teachers' circles on aspects of teachers' mathematical knowledge for teaching. *International Journal for Mathematics Teaching and Learning*. https://eric.ed.gov/?id=EJ1025585

Wickham, H. (2019). *Rvest: Easily harvest (scrape) web pages (R package version 0.3.3)*. https://cran.r-project.org/package=rvest

Williams, G., & Clement, J. (2013, April). *From research to practice: Fostering pre-service science teachers' skills in facilitating effective whole class discussions* [Paper presentation]. Annual conference for the National Association for Research on Science Teaching, Rio Grande, Puerto Rico.

Williams, E. G., & Clement, J. (2014). *Using research on cognitive discussion strategies to support pre-service science teachers' model-based teaching skills* [Paper presentation]. Annual Meeting of the National Association for Research on Science Teaching, Pittsburgh, PA, United States. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.716.978&rep=rep1&type=pdf

Williams, T., Krikorian, J., Singer, J., Rakes, C., & Ross, J. (2019). A high quality educative curriculum in engineering fosters pedagogical growth. *International Journal of Research in Education and Science (IJRES)*, *5*(2), 657–680. https://eric.ed.gov/?id=EJ1223642

Wilson, C. (2016). *Science teachers learning from lesson analysis (STeLLA): High school biology* (Poster). Community for Advancing Discovery Research in Education. https://cadrek12.org/posters/science-teachers-learning-lesson-analysis-stella-high-school-biology

Wilson, C., Stuhlsatz, M., Hvidsten, C., & Stennett, B. (2017). *Examining the impact of lesson-analysis based teacher education and professional development across methods courses, student teaching, and induction* [Oral presentation]. Annual Meeting of the National Association for Research in Science Teaching, San Antonio, TX, United States. https://cadrek12.org/sites/default/files/NARST%20ViSTA%202017.pdf

Wittmann, M. C., Alvarado, C., & Millay, L. (2017). Teacher awareness of problematic facets of meaningful metaphors of energy. *Latin-American Journal of Physics Education*, *11*(2), 2327-1–2327-8. https://dialnet.unirioja.es/servlet/articulo?codigo=6353446

You, H. S., & Kapila, V. (2017, June 25–28). *Effectiveness of professional development: Integration of educational robotics into science and math curricula* [Paper presentation]. 2017 ASEE Annual Conference and Exposition, Columbus, OH, United States. https://peer.asee.org/effectiveness-of-professional-development-integration-of-educational-robotics-into-science-and-math-curricula.pdf

# Appendix A. Review Methodology

This supplemental appendix provides additional details about our review methodology, including the procedures to select the NSF projects, as well as search, select, code, and synthesize their products.

## Project Selection

From NSF's website, we searched for awards meeting the following criteria: (a) had an original award date between January 1, 2011, to December 31, 2015; (b) were tagged with DRK-12 program element code 7645, and (c) were active or completed. This search yielded 428 awards. However, some awards linked to the same project. For instance, a "collaborative research" project will have multiple NSF awards given to separate institutions, though the awards are part of the same project. After removing duplicate awards based on matching titles and abstracts, we identified 376 unique projects. We selected the award date range to focus on projects that were recently completed or are close to completion.

When screening projects for their relevance to PCK, we defined PCK as the knowledge of, reasoning behind, and planning for teaching a specific educational topic or domain (Gess-Newsome et al., 2015, p. 36). Some award abstracts mentioned the term "pedagogical content knowledge" explicitly, whereas others described the concept in other ways (e.g., "teachers' knowledge of students' thinking"). One team member screened the award abstracts for relevance, and another team member reviewed those decisions, discussing any discrepancies with the entire team.

The award abstracts have at least two major limitations: (a) they are brief synopses that likely do not capture the full extent of each project's goals and (b) the project goals may change between the time of award and the time of research. We therefore took two steps to help minimize false negatives and positives (e.g., false negatives are relevant projects missed by our abstract review). First, we shared the list of initially identified projects with our NSF program officer, asking him and other NSF program officers to flag any other PCK-related projects awarded in the eligible time frame. Second, when in doubt, we erred on the side of inclusion, so that we could use the associated products to inform our eligibility decisions. We removed projects if we could not find at least one produced document that was relevant to PCK (see next section). This process yielded 27 PCK-related projects.

## Product Search

We used six sources to identify the publications and resources that the selected projects produced: Web of Science, ERIC, PsycINFO, Google Scholar, Research.gov, and the CADRE

website ([cadrek12.org](cadrek12.org)). This search strategy targeted (a) documents that referenced the numeric award ID and/or (b) documents that project leaders listed on the Research.gov or CADRE websites.

Using the three literature databases (Web of Science, ERIC, and PsycINFO), we searched for the numeric award ID in the funding information search fields (e.g., the grant number field for Web of Science). Using Google Scholar, we searched for documents whose full text contained the numeric award ID and the terms "NSF" or "National Science Foundation." Google Scholar can complement searches of scientific databases by finding relevant gray literature sources (Haddaway et al., 2015). We conducted these searches using the full list of award IDs connected to the 27 PCK-related projects. For instance, a collaborative research project will have multiple award IDs, and we searched for documents containing any of those award IDs. To complement these award ID-based search methods, we developed web scrapers in the *rvest* package in R (Wickham, 2019) to automatically extract citations and other resources (e.g., links to project websites and videos) from the project-specific pages on the Research.gov and CADRE websites. For Research.gov, this search included the public project outcome reports.

We merged the search results from these six different sources using the *revtools* package in R (Westgate, 2018), yielding 333 unique citations after removing duplicates and citations that were not associated with the 27 projects. These citations indexed a diverse set of records and abstracts, including journal articles, conference presentations, book chapters, project websites, project outcome reports, videos hosted on the CADRE website, and other miscellaneous records. The median number of citations per project was eight. In addition, after identifying these 333 citations, we sent emails to each project's PI listing the citations we found, asking the PIs to provide any other products associated with the project.

## Product Selection

For each NSF project, we identified between one to three products that were most closely related to PCK. Product screening occurred in two main phases: (a) identify the products related to PCK and (b) select the one to three products per project that were the most complete and relevant to PCK. We limited the maximum number to three products per project for reasons of practicality (i.e., create a manageable number of products to review) while ensuring representation across projects.

For the first screening phase, we identified documents that reported empirical research (quantitative or qualitative) addressing at least one of the following PCK components in a STEM educational domain:

- **Teachers' Knowledge of Students' Thinking**. Teachers' ideas about students' initial mathematical/science ideas and experiences (including misconceptions), the development

of mathematical/science ideas (including process and sequence), how students express mathematical/science ideas (including demonstration of understanding, questions, and responses), challenging mathematical/science ideas for students, and the appropriate level of mathematical/science understanding.

- **Teachers' Knowledge of Instructional Strategies**. Teachers' knowledge of subject-specific instructional strategies (i.e., strategies specific to teaching mathematics/science, such as fostering scientific inquiry or using the predict-observe-explain pattern) and topic-specific strategies (i.e., strategies specific to teaching particular topics within the domains of mathematics and science, such as energy transfer, fractions, and so on).

- **Teachers' Knowledge of Curriculum**. Teachers' ideas about the scope of mathematics/science (importance of topics and what mathematics/science is worth knowing or teaching), the sequence of mathematics/science (organizing course content for learning), curricular resources available for mathematics/science, and using standards to guide planning and teaching mathematics/science.

- **Teachers' Knowledge of Assessment of Students' Learning**. Teachers' knowledge of strategies for assessing student thinking in mathematics/science and how or when to use assessments.

Screeners considered three questions: (a) Does the study address one of the four components of PCK? (b) Does the study address PCK in a specific STEM educational domain? (c) Is the study an empirical research project (as opposed to literature review or other product)? Documents were retained in screening Phase 1 if the answer was "yes" to all three questions, reducing the number of identified citations from 333 to 201.

For screening Phase 2, we further restricted the corpus by limiting the maximum of products to three per project, for reasons of practicality, as noted earlier. We prioritized products that were

- peer-reviewed (e.g., journal article as opposed to conference poster),

- the most relevant to PCK (e.g., had PCK as its central as opposed to tangential focus), and

- provided the most complete reporting of PCK-related results (e.g., when similar sets of results were reported across multiple products, such as a journal article and a conference paper).

This report's lead author trained junior staff on the screening Phase 1. Training steps included (a) providing example study screening decisions during the initial training phase; (b) listing common reasons for exclusion (e.g., study focused on student, not teacher, cognition); and (c) conducting periodic dual screening checks on the junior screeners' decisions. Screening

phase 2 was, admittedly, more subjective, so the lead author conducted the Phase 2 screening rather than training junior staff.

## Product Structured Coding

We quantitatively coded the products for the presence of key features, such as the component of PCK studied or the research methods used. The "What Was Studied?" section provides results from this coding. As noted in that section, we coded at the product level and then summarized frequencies at the project level (e.g., 14 NSF projects had at least one product that studied the development of teachers' PCK). We summarized at the project level as a meaningful unit of analysis that gave equal weight across projects (rather than weighting toward projects that produced many documents).

We created a sheet using Google Forms with the structured codes and text descriptions for each coding category. For instance, the previous section provides the text descriptions for the PCK components. The lead author trained junior staff on an example set of three study articles, met with them on a weekly basis to address questions about the coding categories, and reviewed their codes to help ensure consistency across coders.

## Synthesis of Empirical Findings

We also summarized the studies' empirical results in three steps. First, we coded lines of text from the results sections using NVivo, categorizing each relevant text section on results about PCK into one of six major lines of empirical research on PCK (see Depaepe et al., 2013, for description of these research lines). Second, we created interpretative bullet point summaries in Microsoft Word for each document, separately by line of PCK research. Third, we further condensed these summaries into Tables S1–S4B4 in Appendix B. We chose a qualitative synthesis approach, rather than a quantitative meta-analysis approach, because the studies varied widely in research methods and were often in qualitative in nature (Thomas & Harden, 2008). For instance, many studies were in-depth qualitative case studies based on interview data, for which extraction of quantitative effect sizes and formal meta-analysis would be inappropriate.

The six lines of PCK research were (a) the nature of PCK, (b) distinguishing PCK from content knowledge, (c) PCK and teaching practice, (d) PCK and student learning, (e) PCK correlates, and (f) PCK development. However, after preliminary data analysis, we dropped two lines of research—distinguishing PCK from content knowledge and PCK correlates—because there were few projects that directly addressed those two topics. Hence, our report instead focuses on four lines of research: (a) nature of PCK, (b) PCK and teaching practice, (c) PCK and student learning, and (d) PCK development.

Teachers' Pedagogical Content Knowledge in Mathematics and Science

## Limitations

Our synthesis focuses only on PCK research funded by NSF's DRK-12 program, meaning it does not cover the entire field of recent STEM education research on PCK. Also, due to limitations in the award abstracts, our synthesis may not cover all recent DRK-12 projects that studied PCK. In addition, the methods rely on publicly available publications and products (or those provided to us by PIs), restricting the observable data about projects to what is reported in these documents. Lastly, the goals, interventions, methods, and outcomes of these projects varied considerably, presenting challenges in coherently synthesizing contributions across projects.

In defense of our synthesis, however, several points are worth noting: (a) the DRK-12 program is a major funder of U.S.-based PK–12 STEM education research; (b) the selected projects are likely representative of the recent DRK-12 portfolio on PCK research, even if some projects might have been missed invertedly; (c) we took extensive effort to find relevant products, including contacting PIs by email; and (d) dividing the results into different lines of empirical research on PCK helped us identify meaningful themes across projects, even if the methods varied.

Teachers' Pedagogical Content Knowledge in Mathematics and Science

# Appendix B. Supplemental Tables

**Table B1. Findings About the Nature of Teachers' Pedagogical Content Knowledge (PCK)**

| NSF award ID and title | PCK domain | Sample size | Type of PCK measure | Key findings related to PCK |
|---|---|---|---|---|
| #1119584: Constructing and Critiquing Arguments in Middle School Science Classrooms: Supporting Teachers With Multimedia Educative Curriculum Materials | Science—scientific argumentation | 34 middle school science teachers for quantitative survey | Quantitative survey with both closed-ended and open-ended response formats | Based on the quantitative survey data (Katsh-Singer et al., 2016), teachers of low-, medium-, and high-socioeconomic status (SES) students did not significantly differ in terms of the overall perceived pedagogical value of argumentation or students' capacity for engaging in argumentation. However, compared with teachers of high-SES students, teachers of low-SES students thought that argumentation was a more important part of their state's science standards ($p < .05$). |
|  |  | 20 middle school science teachers for qualitative interviews | 20- to 30-minute semistructured interviews | Three themes emerged from the interview data (Katsh-Singer et al., 2016):<br>• Teachers widely believe in the value of argumentation activities, but teachers of low-SES students saw different benefits than other teachers.<br>• Teachers hold varying beliefs about students' capacity for argumentation, with many viewing low-SES students, students receiving special education services, and English learners as less capable.<br>• Teachers perceive that pressure from standards and state tests can impact their argumentation instruction, but teachers of high-SES students perceive less of an external influence. |
|  |  | 103 science teachers for PCK instrument development | Multiple-choice quantitative test | Another study (McNeill et al., 2014) developed a multiple-choice assessment of teachers' PCK about scientific argumentation. Key steps of instrument development were (a) conceptualize the domain, (b) design items, (c) pilot-test items with practicing teachers, (d) conduct in-depth cognitive interviews, (e) revise items, (f) solicit advisory board feedback, and (g) finalize items. One key lesson learned about instrument development was that using vignettes is both a strength and weakness; although the real-world context makes the item more authentic, this complexity makes targeting the construct of interest more challenging. |

| NSF award ID and title | PCK domain | Sample size | Type of PCK measure | Key findings related to PCK |
|---|---|---|---|---|
| #1316834: Secondary Science Teaching With English Language and Literacy Acquisition | Science—beliefs about teaching science to English learners | Preservice secondary science teachers (sample size not reported for analysis of interview data) | 30- to 55-minute semistructured interviews conducted before professional development (PD) intervention | Interviews indicated different qualitative patterns, including teachers (a) having narrow conceptions of the role of language and literacy in science teaching; (b) having broadly optimistic expectations of how to teach English learners; (c) feeling underprepared to teach English learners; and (d) reporting some strategies for teaching English learners, such as facilitating productive student talk (Lyon et al., 2016). |
| #1503383: Developing Teachers as Computational Thinkers | Science—infusing computational modeling and thinking into science | 16 middle school science teachers | Qualitative open-ended survey administered after PD intervention and year of teaching | Some teachers reported that students responded well to the student-centered instructional strategies (e.g., reporting that students liked working with a partner; Lee, 2018). Of the instructional strategies taught in the PD program, pair programming was the most commonly used strategy (9 of 16 teachers). However, four teachers indicated how the strategies failed to live up to expectations (e.g., "students struggled with me not helping them right away"). |
| #1503510: Teaching and Learning Algebraic Thinking Across the Middle Grades | Mathematics—using interactive simulations in algebra class | 3 middle school mathematics teachers | Interviews conducted after PD intervention and year of teaching | The three teachers showed three distinct patterns of belief about how the interactive simulations can be used (Findley et al., 2017):<br>• **Simulations as a tradeoff**: They could detract from teacher-student interactions, so they are best used as a supplement at the end of instructional units.<br>• **Simulations as a visual aid**: They promote understanding by visually showing the shapes of graphs and can serve as a starting point for later classroom activities.<br>• **Simulations as an advantage**: They can be a central focus of mathematics lessons to foster student engagement and create opportunities for student discovery. |

| NSF award ID and title | PCK domain | Sample size | Type of PCK measure | Key findings related to PCK |
|---|---|---|---|---|
| #1222777: Assessing, Validating, and Developing Content Knowledge for Teaching Energy | Science—energy topics | 362 high school physics teachers (Etkina et al., 2018) | Quantitative test with constructed-response and multiple-choice item formats (Etkina et al., 2018) | The Content Knowledge for Teaching Energy (CKT-E) instrument was developed through expert teacher item review, pilot testing, and psychometric validation with practicing high school physics teachers. Scoring of constructed-response items showed high interrater reliability (90% or greater), and item response theory models demonstrated sufficient psychometric quality (Etkina et al., 2018). |
|  |  | 25 Grade 6–9 science teachers (Wittmann et al., 2017) | Group discussion activity during PD workshop (Wittmann et al., 2017) | Teacher discussions centered on one item about gravitational potential energy from the CKT-E instrument. Although teachers were aware of several common incorrect student explanations (e.g., energy is "used up"), teachers did not explicitly indicate that they were aware of the common metaphor of energy as a substance-like quantity (Wittmann et al., 2017). |
|  |  | 3 high school physics teachers | Qualitative video-recorded lesson observations (Robertson et al., 2017) | Study developed a qualitative methodology for inferring teachers' CKT-E based on analyzing classroom student-teacher interactions about students' energy ideas. The methodology included processes for selecting relevant classroom episodes of student-teacher interactions, as well as identifying the task of teaching and energy content knowledge addressed. Authors inferred teachers' CKT-E based on how teachers (a) inferred students' model of energy from students' expressions, (b) restated students' ideas, (c) chose an instructional activity to address a misunderstanding, and (d) evaluated student ideas (Robertson et al., 2017). |
| #1503057: Visual Access to Mathematics: Professional Development for Teachers of English Learners | Mathematics—fractions/rational numbers | Middle school mathematics teachers (*n* = 101 for randomized trial; sample size not reported for instrument development) | Quantitative multiple-choice test | The project developed the Mathematical Knowledge for Teaching—Visual Representation (MKT-VR) instrument, consisting of two subscales for content knowledge and PCK for using visual representations, such as diagrams or drawings for teaching ratio and proportional reasoning content. A theoretical framework for the construct of MKT-VR is also described (DePiper & Driscoll, 2018)). A 17-item version of the MKT-VR instrument was used in a randomized trial to assess the efficacy of a PD intervention for teaching with visual representations (Louie & Nikula, 2019). |

| NSF award ID and title | PCK domain | Sample size | Type of PCK measure | Key findings related to PCK |
|---|---|---|---|---|
| #1502778: Building Assessment Items and Instructional Tasks to Build Intercommunity Capacity to Develop Teachers' Mathematical Knowledge for Teaching | Mathematics—basic arithmetic, fractions/rational numbers | 19 elementary teachers identified as most likely having strong MKT (10 or more years of teaching experience) | Think-aloud interview as teachers completed 11 items from the Learning Mathematics for Teaching (LMT) multiple-choice instrument | The pedagogical context—how the assessment items were positioned within a specific instructional scenario—played a critical role in shaping the reasoning leading to correct answers. This context (a) shifts tasks from being disciplinary tasks to being pedagogical mathematics tasks, (b) situates these tasks in contexts requiring doing mathematics with a specific pedagogical purpose, and (c) elicits mathematical reasoning specific to teaching as professional work. The authors argued that attending to the pedagogical context is essential to designing valid MKT assessment items (Hoover & Lai, 2017). |
| | | Teachers included in 190 empirical studies published between 2006 to 2013 (elementary school was the most common grade level) | Systematic review of existing MKT measures, including both quantitative tests and qualitative interviews | The systematic review identified a tension between studies that conceptualize the construct of MKT in broad terms (spanning across all mathematical topics) versus for a specific topic, such as teaching fractions. One key area of progress for the field from 2006 to 2016 was the development of validated instruments. The authors argued for three priorities for future research: (a) finding common ground for engaging in complementary studies, (b) innovating and reflecting on method, and (c) addressing how MKT relates to fluency in teaching and issues of equity/diversity (Hoover et al., 2016). |

| NSF award ID and title | PCK domain | Sample size | Type of PCK measure | Key findings related to PCK |
|---|---|---|---|---|
| #1417838: Knowledge Assets to Support the Science Instruction of Elementary Teachers (ASSET) | Science—small particulate matter and ecosystems | Elementary teachers in Grades 3–5 (*n* = 28 for surveys; *n* = 22 for interviews) | Open-ended online survey and follow-up semistructured qualitative interview | Teachers' responses to the open-ended survey responses (e.g., "Please describe the ideas or misconceptions your students have that make it difficult for them to learn about the particle model of matter") tended to be vague, lacking detail needed to characterize teachers' PCK. The interviews yielded more illuminating responses, though teachers often strayed off topic. The authors argued for a combined measurement approach in which survey responses give an overview of the teachers' PCK and then interviewers could probe for more detailed responses (Smith et al., 2017).

The results suggested that the elementary teachers have limited PCK about small particulate matter as an explanatory model. The authors attributed this lack of PCK as a lack of opportunity or need to teach about the topic in elementary school. Only the most recent national standards (i.e., the Next Generation Science Standards) have included the small particle model in elementary science standards (Smith et al., 2017).

An overview chapter also described a similar survey-interview approach for eliciting elementary teachers' PCK about ecosystems (Smith et al., 2018). Lastly, a systematic review surveyed the literature for studies on student understanding and topic-specific instructional strategies about small particulate matter and ecosystems; the aim of the review was to synthesize such knowledge to make it more practically useful for teachers and support their PCK development (Smith et al., 2016). |

| NSF award ID and title | PCK domain | Sample size | Type of PCK measure | Key findings related to PCK |
|---|---|---|---|---|
| #1316653: Theorizing and Advancing Teachers' Responsive Decision Making in the Domain of Rational Numbers | Mathematics—fractions/rational numbers | 18 elementary teachers in Grades 3–5 (mixed years of experience) | Open-ended survey; teachers were asked to provide five different valid strategies for how students might solve a specific fractions problem. | The study's analysis focused on categorizing teachers into three different levels of PCK (robust, limited, or lack of PCK) based on how teachers' anticipated strategies were consistent with children's typical strategies. The definition of "typical" strategies was based on prior research on children's thinking about fractions. Teachers in the lowest knowledge category (i.e., lack of evidence of PCK) still often generated possible strategies, but the researchers judged those strategies to be atypical based on prior research on student understanding about fractions (Krause et al., 2016). |
| | | 3 elementary teachers in Grades 3–5 | Lesson observations, including videos of 2 days of classroom instruction and one-on-one interactions with students | Another study proposed a theoretical model in which teachers' knowledge of children's mathematical thinking is a component of responsive teaching that can provide opportunities for children to advance their thinking. Video observations illustrated how expert teachers used this knowledge to notice and advance children's mathematical thinking in classroom and one-on-one interactions (Jacobs et al., 2015). |
| #1503206: Student Adaptive Pedagogy for Elementary Teachers | Mathematics—basic arithmetic, fractions/rational numbers | 2 elementary teachers in Grades 3–5 | In-depth qualitative case studies with interviews, observations at PD sessions, and lesson observations | This in-depth qualitative study illustrated how two teachers' explanations of students' mathematical reasoning shifted across time from being based on teachers' own reasoning (first-order model) toward a more explicit recognition of how students' reasoning differs from teachers' reasoning. Manifestations of this shift include (a) juxtaposing teachers' thought processes with students', (b) deepening ability to think about students' mathematical reasoning, (c) enhanced skills to depict students' reasoning, and (d) growing mindfulness and intention to use students' reasoning to guide instructional decisions (Hodkowski, 2018). |

| NSF award ID and title | PCK domain | Sample size | Type of PCK measure | Key findings related to PCK |
|---|---|---|---|---|
| #1316571: Investigating Simulations of Teaching Practice: Assessing Readiness to Teach Elementary Mathematics | Mathematics—basic arithmetic | 47 preservice elementary teachers (Shaughnessy & Boerst, 2017) | Qualitative interview in a standardized clinical teaching simulation (Shaughnessy & Boerst, 2017) | This study presented novice teachers with a student's work on a specific addition problem, after which teachers had 5 minutes to interact with a standardized "student" to diagnose the simulated student's reasoning. Researchers developed a checklist of core desired practices, including (a) eliciting the student's process, (b) probing understanding of key mathematical ideas, (c) attending to the student's ideas, and (d) deploying other moves that support learning (Shaughnessy & Boerst, 2017). |
| | | 3 preservice elementary teachers (Shaughnessy et al., 2018) | Clinical simulation interviews and lesson observations | A subsequent study contrasted this clinical simulation assessment with a field assessment in which preservice teachers interacted with real elementary students. The study focused on three specific teachers who were chosen to illustrate individual differences. The results showed the simulation and field assessments each offered unique affordances and challenges. Although the field assessment has greater authenticity of representing real-world teaching practice, the variability of students and situationally specific demands presents interpretational challenges; the simulation assessment addresses this issue with a standardized student, but the simulation has less real-world authenticity (Shaughnessy et al., 2018). |
| #1417769: Teaching STEM With Robotics: Design, Development, and Testing of a Research-Based Professional Development Program for Teachers | Science, mathematics, and engineering—robotics instruction | 20 middle school teachers | Quantitative survey about teachers' PCK and the perceived teacher knowledge requirements for delivering robotics-focused instruction | Teachers had the lowest self-efficacy for the technical knowledge (TK) component (e.g., "I have the technical skills that I need to teach my robotics-based lesson") but had higher self-efficacy for the PCK component (e.g., "I can select effective teaching approaches to guide students' thinking and learning in mathematics/science in robotics-focused lessons"). The TK component was also rated as the most important teacher prerequisite for teaching robotics-focused lessons (Rahman et al., 2017). |

Teachers' Pedagogical Content Knowledge in Mathematics and Science

| NSF award ID and title | PCK domain | Sample size | Type of PCK measure | Key findings related to PCK |
|---|---|---|---|---|
| #1316601: Engineering for All (EfA) | Engineering design | 21 middle school engineering teachers | Engineering design teaching portfolios (qualitatively analyzed using teaching design rubrics to infer teachers' PCK) | The teaching portfolios included (a) instructional logs after each lesson/unit, (b) student work, (c) instructional video clips, and (d) written reflections on needed curricular revisions. The qualitative data illustrated the varying levels of understanding of how to teach engineering design. The ways that teachers adapted the curriculum or not was one indicator of their PCK. Videos showed that some teachers addressed students' misconceptions, whereas other teachers did not elicit students' thinking and therefore could not address students' misconceptions (Crismond & Lomask, 2016). These empirical findings and other studies in the broader literature informed a theoretical framework and literature review on characterizing teachers' PCK for fostering design-based learning and instruction (Crismond & Adams, 2012). |
| #1118772: The Science and Mathematics Simulated Interaction Model (SIM) | Mathematics—geometry | 8 preservice mathematics teachers (Dotger et al., 2015) | Clinical simulation with standardized data (interview data qualitatively analyzed) | For the mathematics study (Dotger et al., 2015), the standardized students were given a four-page protocol that detailed a set of triggers—questions, statements, concerns, or responses by the students—to be issued during the simulation. Preservice teachers' interactions with these students were coded across four primary constructs: (a) diagnosis, (b) explanation, (c) mathematics repertoire, and (d) teaching/instructional repertoire. The study authors emphasized how the qualitative data could offer opportunities for diagnosing and fostering preservice teachers' PCK and skills for interacting one-on-one with students. The science study (Dotger et al., 2018) applied a similar approach but in the context of student misconceptions about natural selection. |
| | Science—biology/life science, natural selection | 13 preservice science teachers (Dotger et al., 2018) | | |

| NSF award ID and title | PCK domain | Sample size | Type of PCK measure | Key findings related to PCK |
|---|---|---|---|---|
| #1316736: Improving Formative Assessment Practices: Using Learning Trajectories to Develop Resources That Support Teacher Instructional Practice and Student Learning in CMP2 | Mathematics—teacher ideas about learning trajectories | 10 middle school mathematics teachers | Qualitative interview data about teachers' instructional and assessment practices | Teachers' descriptions of student learning sequences and common learning obstacles differed in terms of three grain sizes: (a) grade level, (b) instructional unit level, and (c) mathematical topic level. Teachers generally talked about student challenges at a large grain size (e.g., students struggle with fractions) and seldom at a finer grain sizes. Teachers also described the instructional challenges in teaching students with gaps in prior knowledge, such as a lack of proficiency in previously taught content (Castro Superfine & Li, 2017). |
| #1118894: PROJECT ATOMS: Accomplished Elementary Teachers of Mathematics and Science | Mathematics—whole numbers and rational numbers; science—life science and physical science | 55 preservice teachers enrolled in an elementary education teacher preparation program | Quantitative test with correct/incorrect answers and interviews | Survey data found that preservice teachers with higher PCK also had higher teacher self-efficacy. Interview data identified how teachers' college coursework and college field experiences influenced their PCK beliefs (Thomson et al., 2016). |
| #1222709: Identifying Science Teaching Strategies for Promoting Reasoned Discussions of Concepts and Simulations | Science—electricity, model-based learning and leading classroom discussions | Preservice teachers (unknown sample size) | Interviews and open-ended surveys | The average preservice teacher thought that whole-class discussions should spend most of their time on student-generated ideas, including student evaluation of those ideas. These teachers thought that effective strategies for sustaining whole-class discussion include giving everyone a chance to talk, using good questions, redirecting, and probing. |

## Table B2. Findings About the Development of Teachers' PCK

| NSF award ID and title | PCK domain | Sample size | Research design and/or intervention | Teacher PCK measure | Likely WWC rating | Key findings |
|---|---|---|---|---|---|---|
| [#1417895: Preparing Urban Middle Grades Mathematics Teachers to Teach Argumentation](#) | Mathematics—similarity, coordinate geometry | 31 middle school mathematics teachers | • Randomized controlled trial comparing alternative versions of PD program (treatment and comparison). <br>• Both versions focused on MKT in Week 1. But in Week 2, only the treatment version had a bridging component connecting MKT to classroom practice through improvisational teaching games and planning through visualization. | MKT (quantitative multiple-choice measure) | • Does Not Meet WWC Group Design Standards <br>• Insufficient information reported to assess attrition or baseline equivalence | Both treatment and comparison teachers' MKT significantly increased from before to after the PD workshop. The report did not indicate if this change was greater for treatment than for comparison teachers (Knudsen et al., 2015). |
| [#1503206: Student-Adaptive Pedagogy for Elementary Teachers](#) | Mathematics—multiplicative and fractional reasoning | Two elementary teachers in Grades 3–5 | • Longitudinal qualitative study <br>• Intervention was three, job-embedded PD experiences with a focus on building teachers' PCK about students' multiplicative reasoning (110 hours) | In-depth qualitative case studies with interviews, observations at PD sessions, and lesson observations | • Ineligible for review under WWC Group Design Standards <br>• No comparison group | This in-depth qualitative study illustrated how two teachers' explanations of students' mathematical reasoning shifted across time from being based on teachers' own reasoning (first-order model) toward a more explicit recognition of how students' reasoning differs from teachers' (see Table B1 for more details; Hodkowski, 2018). |

| NSF award ID and title | PCK domain | Sample size | Research design and/or intervention | Teacher PCK measure | Likely WWC rating | Key findings |
|---|---|---|---|---|---|---|
| #1118643: Energy: A Multidisciplinary Approach for Teachers (EMAT): Designing and Studying a Multidisciplinary, Online Course for High School Teachers | Science—energy | 47 in-service high school teachers | • Pre-post longitudinal design<br>• Intervention was a 10-week, online PD summer course (120 hours) focused on energy topics using video-based analysis of classroom instruction. | Teacher ability to analyze video clips of science teaching and learning (constructed-response format scored with a rubric) | • Ineligible for review under WWC Group Design Standards<br>• No comparison group | Teachers' ability to analyze videos significantly increased after treatment (effect size [ES] = 1.38, $p < .001$; Kowalski, 2016; Kowalski et al., 2018). |
| #1220635: Videocases for Science Teaching Analysis Plus (ViSTA Plus) | Science—general/ specific content area not reported | 45 preservice teachers from two universities (based on sample sizes at student teaching) | • Quasi-experimental with business-as-usual comparison<br>• Intervention was a semester-long methods course for preservice elementary teachers using video-based analysis of classroom instruction. | Teacher ability to analyze video clips of science teaching and learning (constructed-response format scored with a rubric) | • Does Not Meet WWC Group Design Standards<br>• Insufficient information reported to assess baseline equivalence for the posttreatment analytic sample | PCK was higher for treatment teachers than comparison teachers, both directly after the methods course (ES = 1.68, $p < .001$) and during student teaching (ES = 0.74, $p = .012$; Roth et al., 2018; Wilson et al., 2017). |

| NSF award ID and title | PCK domain | Sample size | Research design and/or intervention | Teacher PCK measure | Likely WWC rating | Key findings |
|---|---|---|---|---|---|---|
| #1118894: PROJECT ATOMS: Accomplished Elementary Teachers of Mathematics and Science | Science—general/specific content area not reported | 12 preservice elementary school teachers for survey data (*n* = 3 for interview data) | • Longitudinal design (four time points)<br><br>• Teacher preparation program; STEM education researchers taught the science methods courses. | • Quantitative survey on teachers' beliefs about learning-theory-aligned science instruction (e.g., "teachers should ask students to support their conclusions with evidence").<br><br>• Qualitative interviews about teachers' visions for science instruction | • Ineligible for review under WWC Group Design Standards<br><br>• No comparison group | Teachers were grouped into three patterns of change: (a) increasing, (b) decreasing, or (c) stable beliefs about learning-theory-aligned science instruction.<br><br>One teacher from each group was selected for in-depth case analysis using qualitative interview data (Carrier et al., 2018). |
| #1119163: Implementing the Mathematical Practice Standards: Enhancing Teachers' Ability to Support the Common Core State Standards | Mathematics—geometry and algebra | 88 mathematics teachers in Grades 5–10 in the treatment group | • Quasi-experimental, pre-post design with business-as-usual comparison<br><br>• Ten 2-hour PD sessions to increase understanding of the Standards for Mathematical Practice (SMP) described in the Common Core Mathematics Standards | Quantitative test about teachers' (a) awareness of the SMP, (b) readiness to use the SMP, (c) knowledge of the SMP, and (d) identification of the SMP in examples of mathematical thinking. | • Does Not Meet WWC Group Design Standards<br><br>• Insufficient information reported to assess baseline equivalence | Relative to comparison teachers, treatment teachers scored higher on all four subscales after the PD, controlling for pretreatment scores (Goldenberg et al., n.d.). |
| #1119342: Investigating the Impact of Math Teachers' Circles on Mathematical Knowledge for Teaching and Classroom Practice | Mathematics | 50 middle school mathematics teachers | • Pre-post longitudinal design<br><br>• A 1-week summer PD program, facilitated by professional mathematicians, focused on developing teachers' mathematical problem-solving skills | MKT assessment (Number Concepts and Geometry subscales) | • Ineligible for review under WWC Group Design Standards<br><br>• No comparison group | Teachers' MKT significantly increased across time for the Number Concepts subscale but not for the Geometry subscale. The interaction with the outcome domain, however, was also not significant (White et al., 2013). |

| NSF award ID and title | PCK domain | Sample size | Research design and/or intervention | Teacher PCK measure | Likely WWC rating | Key findings |
|---|---|---|---|---|---|---|
| #1503057: Visual Access to Mathematics: Professional Development for Teachers of English Learners | Mathematics— fractions/ rational numbers | 101 middle school mathematics teachers | • Cluster randomized controlled trial<br><br>• 1-year 60-hour blended learning PD course on using visual representations in mathematics instruction | Two quantitative multiple-choice tests: (a) 28 MKT items from the Ratio and Proportional Reasoning subscales and (b) 17 items from a researcher-developed test on MKT with visual representations | • Does Not Meet WWC Group Design Standards<br><br>• Insufficient information reported to assess attrition or baseline equivalence | Treatment effect was not statistically significant for either PCK measure. However, relative to control teachers, treatment teachers reported significantly greater self-efficacy for using visual representations to teach mathematics (DePiper & Driscoll, 2018; Louie & Nikula, 2019). |
| #1417769: Teaching STEM With Robotics: Design, Development, and Testing of a Research-Based Professional Development Program for Teachers | Science— robotics | 20 middle school mathematics and science teachers | • Pre-post longitudinal design<br>• 3-week summer PD program (120 contact hours) about using robotics kits to teach standards-aligned science and mathematics curricula | Questionnaire about PCK self-efficacy for using robotics to teach science concepts | • Ineligible for review under WWC Group Design Standards<br><br>• No comparison group | Teachers' PCK self-efficacy significantly increased after the professional development workshop (ES = 0.68, $p < .001$; You & Kapilla, 2017). A qualitative study with 4 teachers indicated their perceptions of different student challenges with learning from the curriculum (Brill et al., 2016). |

| NSF award ID and title | PCK domain | Sample size | Research design and/or intervention | Teacher PCK measure | Likely WWC rating | Key findings |
|---|---|---|---|---|---|---|
| #1316736: Improving Formative Assessment Practices: Using Learning Trajectories to Develop Resources That Support Teacher Instructional Practice and Student Learning in CMP2 | Mathematics—teacher ideas about learning trajectories | 10 middle school mathematics teachers | • Longitudinal design (teachers interviewed at least 3 times during the first year of the PD program)<br><br>• The PD program included formative assessment resources, an online platform, and sessions about learning trajectories of algebra concepts. | Qualitative interview data about teachers' instructional and assessment practices | • Ineligible for review under WWC Group Design Standards<br><br>• No comparison group | Results about change across time in teacher PCK were not reported in the conference abstracts (Bragelman et al., 2017; Castro Superfine & Li, 2017). Other results indicated challenges in delivering the online professional development, such as reflection fatigue and inequitable participation (Bragelman et al., 2017). |
| #1222709: Identifying Science Teaching Strategies for Promoting Reasoned Discussions of Concepts and Simulations | Science—perceptions about whole-classroom discussions in science | 17 preservice teachers (intended grade level not reported) | • Pre-post longitudinal design<br><br>• 8-week course unit (32 hours) for preservice teachers about leading model-based whole-classroom discussions in science classes | Quantitative survey about the perceived importance of whole-classroom discussions in science classes | • Ineligible for review under WWC Group Design Standards<br><br>• No comparison group | Teachers significantly increased in their beliefs about the usefulness of whole-classroom discussions in science classes, in addition to the importance of fostering specific science inquiry activities (e.g., have students make predictions) within those whole-class discussions.<br><br>Also, the percentage of classroom time that teachers thought should go to whole-class discussions increased from 18% to 36%, whereas the percentage of time spent on lectures decreased (Williams & Clement, 2014). |

| NSF award ID and title | PCK domain | Sample size | Research design and/or intervention | Teacher PCK measure | Likely WWC rating | Key findings |
|---|---|---|---|---|---|---|
| #1317068: Improving Competency in Elementary Science Teaching | Science—teaching scientific practices | 90 elementary science teachers | • Single-group observational study<br>• Planning and reflection tool, paired with a video database of classroom episodes and professional development | No explicit indication of PCK being measured | • Ineligible for review under WWC Group Design Standards<br>• No comparison group | The project aimed to build professional development resources for fostering teachers' PCK, but the project outcomes report did not indicate if this outcome was evaluated (Darling-Hammond, n.d.). |
| #1503280: Science Teachers Learning from Lesson Analysis (STeLLA): High School Biology | Science—high school biology | High school biology teachers (unknown sample size) | • Cohort control design in which comparisons are made with an earlier school year<br>• Professional development focused on video-based analysis of science lessons | Quantitative test with correct/incorrect answers | • Does Not Meet WWC Group Design Standards<br>• School year is a confounding factor in a cohort control design. | Poster described the study design but not results because the project was in the development phase at the time of the presentation (Wilson, 2016). |
| #1503399: An Efficacy Study of the Learning and Teaching Geometry Professional Development Materials: Examining Impact and Context-Based Adaptations | Mathematics—geometry | 103 middle school and high school mathematics teachers | • School-level randomized controlled trial<br>• 9-day PD program | Multiple quantitative tests, including the Diagnostic Science Assessments for Middle School Teachers and a separate set of PCK assessments tailored to the intervention | • Meets WWC Group Design Standards Without Reservations<br>• Low-attrition cluster-level randomized controlled trial | Intervention teachers increased in all knowledge measures pre to post, but only significantly more so than control teachers for the PCK measures tailored to the intervention (Jacobs et al., 2019).<br>Another study provided an in-depth description of the PD curriculum, focusing on video-based analysis of instructional episodes (Seago et al., 2018). |

## Table B3. Findings About the Relationship of Teachers' PCK With Teaching Practice and Quality

| NSF award ID and title | PCK domain | Sample size | Research design and/or intervention | Teacher PCK measure | Teaching practice measure | Key findings |
|---|---|---|---|---|---|---|
| #1417895: Preparing Urban Middle Grades Mathematics Teachers to Teach Argumentation | Mathematics—similarity, coordinate geometry | 31 middle school mathematics teachers | • Randomized controlled trial comparing alternative versions of PD program (treatment and comparison)<br>• Both versions focused on MKT in Week 1. But in Week 2, only the treatment version had a bridging component connecting MKT to concrete plans for classroom practice. | MKT (quantitative multiple-choice measure) | Videotaped classroom practice observations | Treatment teachers more often engaged in teaching moves that were coded as supporting argumentation, such as asking open-ended questions and encouraging the participation of multiple students (Knudsen et al., 2015).<br><br>The study also reported that teachers' MKT correlated with argumentative talk in the classroom, though it is unclear if this result meant that teachers' MKT correlated with teaching moves versus student talk. Presumably, a correlation with student talk should be mediated by teaching practices. |

Teachers' Pedagogical Content Knowledge in Mathematics and Science

| NSF award ID and title | PCK domain | Sample size | Research design and/or intervention | Teacher PCK measure | Teaching practice measure | Key findings |
|---|---|---|---|---|---|---|
| #1118643: Energy: A Multidisciplinary Approach for Teachers (EMAT): Designing and Studying a Multidisciplinary, Online Course for High School Teachers | Science—energy topics | 47 in-service high school teachers | • Pre-post comparison at the teacher level<br>• Intervention was a 10-week, online, video-based PD summer course (120 hours) focused on energy topics. | Teacher ability to analyze video clips of science teaching and learning (constructed-response format scored with a rubric) | Videotaped classroom practice observations (coded for use of instructional strategies taught in the PD program) | Teachers' use of the taught instructional strategies tended to increase across time, but the change was nonsignificant ($p$ = .063).<br>Preintervention teacher practice did not predict postintervention teacher practice, but it is unclear from the report how MKT correlated with teacher practice (Kowalski et al., 2018). |
| #1503510: Teaching and Learning Algebraic Thinking Across the Middle Grades | Mathematics—using interactive simulations in algebra class | 3 middle school mathematics teachers | Observational study conducted after PD intervention and a year of teaching using interactive simulations | Qualitative interviews | Qualitative analysis of lesson plans | Individual differences in teachers' use of interactive simulations in the classroom (i.e., teacher practice) closely aligned with teachers' beliefs (i.e., PCK) about the affordances and drawbacks of using interactive simulations to teach mathematics concepts (Findley et al., 2017). |
| #1502778: Building Assessment Items and Instructional Tasks to Build Intercommunity Capacity to Develop Teachers' Mathematical Knowledge for Teaching | Mathematics—several topics | Literature review of 12 empirical studies of how MKT relates to teaching practice (most studies had fewer than 10 teachers) | Literature review of mostly qualitative, observational studies | Mostly qualitative approaches, such as interviews | Mostly qualitative approaches, such as classroom observations | The review found that prior research has often focused on identifying the knowledge that teachers need and possess, but the authors argued it would be more useful to focus on detailing the work of doing mathematics teaching (Mosvold & Hoover, 2017). |

| NSF award ID and title | PCK domain | Sample size | Research design and/or intervention | Teacher PCK measure | Teaching practice measure | Key findings |
|---|---|---|---|---|---|---|
| #1220635: Videocases for Science Teaching Analysis Plus (ViSTA Plus) | Science—specific content area not reported | 45 preservice teachers from two universities (based on sample sizes at student teaching) | • Quasi-experimental with business-as-usual comparison<br>• Intervention was a semester-long methods course for preservice elementary teachers using video-based analysis of classroom instruction. | Teacher ability to analyze video clips of science teaching and learning (constructed-response format scored with a rubric) | Video-recorded lessons during preservice teachers' field student teaching (the process for coding and analyzing these videos was not reported in Wilson et al., 2017) | Teacher practice scores were higher for treatment teachers than comparison teachers (ES = 2.05, $p < .001$). Baseline equivalence was not reported for this quasi-experimental study (Wilson et al., 2017). |

## Table B4. Findings About the Impact of Teachers' PCK on Student Learning Outcomes

| NSF award ID and title | PCK domain | Sample size | Research design and/or intervention | Teacher PCK measure | Student outcomes | Key findings |
|---|---|---|---|---|---|---|
| #1417895: Preparing Urban Middle Grades Mathematics Teachers to Teach Argumentation | Mathematics—similarity, coordinate geometry | 31 middle school mathematics teachers for randomized controlled trial (unknown number of students) | • Randomized controlled trial comparing alternative versions of PD program (treatment & comparison)<br>• Both versions focused on MKT in Week 1. But in Week 2, only the treatment version had a bridging component connecting MKT to classroom practice through improvisational teaching games and planning through visualization. | MKT (quantitative multiple-choice measure) | Student argumentative talk (classroom observations) | The frequency of student argumentative talk was higher in the treatment group. The average length of student argumentative talk was related to teachers' MKT (Knudsen et al., 2015). |
| #1503206: Student-Adaptive Pedagogy for Elementary Teachers | Mathematics—multiplicative and fractional reasoning | • 7 teachers for quasi-experimental comparison in Grades 3–5<br>• 205 elementary school students for quasi-experimental comparison | • Two research designs were used: (a) quasi-experimental comparison at one school and (b) pre-post gains at two schools<br>• Intervention was three, job-embedded PD experiences with a focus on building teachers' PCK about students' multiplicative reasoning (110 hours). | Teacher PCK was not measured in this study on student outcomes (Tzur et al., 2018). | Students' multiplicative reasoning (5-item written quantitative assessment) | • For the quasi-experimental study, pre-post gains were larger for students of intervention teachers versus comparison teachers (Tzur et al., 2018).<br>• For the pre-post study, students made significant gains in their multiplicative reasoning (Tzur et al., 2018). |

| NSF award ID and title | PCK domain | Sample size | Research design and/or intervention | Teacher PCK measure | Student outcomes | Key findings |
|---|---|---|---|---|---|---|
| #1118643: Energy: A Multidisciplinary Approach for Teachers (EMAT): Designing and Studying a Multidisciplinary, Online Course for High School Teachers | Science—energy topics | • 47 in-service high school teachers<br>• 2,462 high school students | • Quasi-experimental comparison using students from a prior cohort (same teachers but students from the preintervention year)<br>• Intervention was a 10-week, online PD summer course (120 hours) focused on energy topics using video-based analysis of classroom instruction. | Teacher ability to analyze video clips of science teaching and learning (constructed-response format scored with a rubric) | Student content knowledge about energy topics (35-item multiple-choice quantitative test) | • The estimated effect on student knowledge was 0.13 standard deviations and nonsignificant (in contrast to a 0.68 standard deviation significant effect in a prior randomized controlled trial with a face-to-face implementation of the PD).<br>• Teachers' ability to analyze videos also did not significantly predict student outcomes (Kowalski, 2016; Kowalski et al., 2018). |
| #1220635: Videocases for Science Teaching Analysis Plus (ViSTA Plus) | Science—specific content area not reported | 45 preservice teachers from two universities (based on sample sizes at student teaching) | • Quasi-experimental with business-as-usual comparison<br>• Intervention was a semester-long methods course for preservice elementary teachers using video-based analysis of classroom instruction. | Teacher ability to analyze video clips of science teaching and learning (constructed-response format scored with a rubric) | Student science achievement assessed via curriculum unit tests during preservice teachers' field training | Student achievement was higher for treatment versus comparison teachers (ES = 0.38, p = .01). Correlations between teacher PCK and student outcome measures were not reported (Roth et al., 2018; Wilson et al., 2017). |
| #1118894: PROJECT ATOMS: Accomplished Elementary Teachers of Mathematics and Science | Mathematics—number and operations | 118 novice elementary teachers | • Correlational design (intervention group only; no baseline data)<br>• One-day summer PD session before teachers' second year of teaching | MKT (Number and Operations subscale) | Videos of teachers' instruction were coded for the level of student argumentative talk using the Mathematics Scan Observation measure. | Teachers' MKT correlated with the level of student explanation and justification but not the overall level of mathematical discourse community (Lee & Walkowiak, 2016). |

## About the American Institutes for Research

Established in 1946, with headquarters in Arlington, Virginia, the American Institutes for Research® (AIR®) is a nonpartisan, not-for-profit organization that conducts behavioral and social science research and delivers technical assistance to solve some of the most urgent challenges in the U.S. and around the world. We advance evidence in the areas of education, health, the workforce, human services, and international development to create a better, more equitable world. For more information, visit AIR.ORG.