Center on
GREAT TEACHERS & LEADERS
at the American Institutes for Research®

How Learning Forward's Standards for Professional Learning Are Associated With Teacher Instruction and Student Achievement: A Meta-Analysis

Rachel Garrett, Qi Zhang, Martyna Citkowicz, and Lauren Burr

December 2021

Introduction

The periodic revision of Learning Forward's Standards for Professional Learning provides the opportunity to engage with research related to teaching and student learning. In 2020, Learning Forward began work to revise the Standards for Professional Learning and will release the updated standards in 2022. As part of the standards revision effort, Learning Forward sought to leverage insights from research that has been published since the release of the standards in 2011. This work has included drawing on more recent research published since the 2011 Standards to understand the relationship between teacher professional learning programs that embody features of the standards with educational outcomes among teachers and students. To support this effort, the Center on Great Leaders and Teachers at the American Institutes for Research conducted a systematic review and meta-analysis to understand the relationships between the standards and teacher and student outcomes. While the standards are relevant for all types of educators, in this study, we focused on classroom teachers as a key group that had a sufficient research base available to study. The systematic review gathers the relevant literature using a formal and structured approach, and the meta-analysis employs a statistical procedure to synthesize information across studies. This two-stage process allows for a rigorous examination of the consistency of the relationships between program standards and desired outcomes across studies.

The present meta-analysis serves to shed light on the context in which the standards sit within existing research in the last decade. Its purpose is to formally analyze available evidence from a specific type of research—randomized field trials of teacher professional learning programs—that ties features of the standards to teacher instruction and student learning. Our analyses are grounded in an initial examination of the 2011 Standards, along with additional areas of professional learning approaches identified by Learning Forward as important to the field. Once the draft 2022 revised standards were available, we turned to examining how the new standards were reflected in the research and associated with teacher and student outcomes. In addition, we conducted one of the first mediation meta-analyses in this area. This approach allowed us to empirically quantify how changes in teacher instruction were associated with changes in student achievement—a key tenant undergirding theories of teacher professional learning that have had minimal analysis in research to date.

The findings from this work provide Learning Forward with information about the research base of randomized field studies supporting the 2022 Standards. Drawing on rigorous empirical research through a meta-analytic approach also helps identify areas for future research by highlighting what remains unknown about program outcomes. Consequently, the findings of this meta-analysis can inform Learning Forward's research agenda that will direct the field.

The Professional Standards and Areas for Deeper Exploration

The 2011 Standards for Professional Learning encompassed seven standards to define high-quality and aspirational ways to implement professional learning for all educators.

The Learning Forward 2011 Standards for Professional Learning

Professional learning that increases educator effectiveness and results for all students...

- occurs within **learning communities** committed to continuous improvement, collective responsibility, and goal alignment;
- requires skillful **leaders** who develop capacity, advocate, and create support systems for professional learning;
- requires prioritizing, monitoring, and coordinating **resources** for educator learning;
- uses a variety of sources and types of student, educator, and system **data** to plan, assess, and evaluate professional learning;
- integrates theories, research, and models of **learning designs** to achieve its intended outcomes;
- applies research on change and sustains support for **implementation** of professional learning for long-term change; and
- aligns its **outcomes** with educator performance and student curriculum standards.

Since the setting of the 2011 Standards, a growing body of research and implementation in the field surfaced additional areas of interest salient to teacher professional learning. These areas have emerged as Learning Forward has engaged in ongoing discussion with practitioners and researchers. For example, experts in student learning standards have helped to map out alignment of student learning standards with the Learning Forward professional learning standards. This helped identify areas in which the Learning Forward standards had opportunities to become more robust with respect to the teacher role in the student standards. Further, Learning Forward has engaged with states across the country as they adopt or adapt their standards. This interaction has surfaced additional needed components over time, such as states adding an equity standard (Herpin, 2019). Through these discussions, implementation of the 2011 Standards, and recent research, Learning Forward identified the following eight Areas for Deeper Exploration to position and inform the 2022 standards revisions.

Areas for Deeper Exploration	
	**Equity and Cultural Competence |** Equity (e.g., addressing achievement gaps, supporting underserved populations) is foundational to professional learning content, delivery, and outcomes.
	**Learning Systems |** Learning systems include a shared vision for the role of professional learning in school improvement, dedicated resources for sustained learning, and alignment with all school and system functions.
	**Content Expertise |** Content expertise enables teachers to understand and apply strategies to ensure that students can meet and exceed grade-level and course expectations.
	**Instructional Materials |** Teachers become experts in the selection and use of instructional materials that anchor teaching in evidence-based practices for student learning.
	**Personalized Learning for Students |** Teachers learn about and apply proven and promising approaches to personalized learning that support models of individualized instruction for diverse learners.
	**Personalized Learning for Teachers |** Teachers engage in personalized learning through individualized coaching, micro-credentialing, technology, and related approaches that support individualized pathways to professional growth.
	**Social and Emotional Learning for Students |** Content focuses on the principles of social and emotional learning to ensure that teachers are equipped to create safe, healthy classroom cultures that enable students to thrive.
	**Technology and Learning Innovations |** Teachers incorporate relevant technology and learning platforms and innovations that have the potential to impact teacher and student learning, both in person and virtually.

In the current revision, the 2022 Professional Learning Standards have evolved from seven to 11 standards, organized into three frames that recognize (1) the conditions necessary for professional learning to succeed, (2) high-quality processes for creating effective professional learning experiences, and (3) critical areas of content for educators to focus on as they continue to grow their expertise.

The Learning Forward 2022 Standards for Professional Learning

Professional learning results in equity and excellent outcomes for all students when educators . . .

Conditions for Success

Equity Foundations | Establish expectations for equity, create structures to ensure all staff members have access to learning, and sustain a culture of support for all staff.

Culture of Collaborative Inquiry | Commit to continuous improvement, build collaboration skills and capacity, and share responsibility for improving learning for all students.

Leadership | Establish a compelling and inclusive vision for professional learning, ensure a coherent system of support to build individual and collective capacity, and advocate for professional learning by sharing the importance and evidence of impact of professional learning with others.

Resources | Allocate resources for professional learning, prioritize equity in their resource decisions, and monitor the use and impact of resource investments.

Transformational Processes

Equity Drivers | Identify and address their own biases, collaborate with diverse colleagues, and cultivate beliefs, knowledge, and behaviors that support each learner.

Evidence | Consider evidence and data from multiple sources, use data to plan and monitor learning, and assess impact of professional learning on educators and students.

Learning Designs | Set relevant and contextualized learning goals, ground their work in research and theories about learning, and implement evidence-based learning designs.

Implementation | Understand and apply research on change management theory, engage in feedback processes, and implement and sustain professional learning.

Rigorous Content for Each Learner

Equity Practices | Understand their students' historical and societal contexts, embrace student assets through instruction, and foster relationships with students, families, and communities.

Curriculum, Assessment, and Instruction | Provide high-quality curriculum and instructional materials for students, assess student learning, and understand curriculum and implement through instruction.

Professional Expertise | Develop the expertise essential to their roles, apply standards and research to their work, and prioritize coherence and alignment in their learning.

Study Approach

Research Questions

In this study, we addressed the following set of research questions:

- What is the evidence of how PD programs that contain elements aligned to the 2011 Standards or Areas for Deeper Exploration are associated with changes in teacher instruction and student learning?
- What is the evidence of how PD programs that contain elements aligned to the 2022 Standards are associated with changes in teacher instruction and student learning?
- What does the evidence indicate about the extent to which changes in instruction mediate changes in student learning?

Systematic Literature Search

We first conducted a systematic literature search to find studies that involved randomized field trials of teacher professional learning programs that were published between 2010 and 2020. To be selected for the meta-analysis, studies needed to (1) include inservice teachers in Grades K–12, (2) examine the impacts of professional learning on instruction measured through classroom observation,[1] and (3) have enough information to compute effect sizes (see Study Coding section). We searched for studies in multiple research literature databases and, after screening the identified studies for our criteria, we identified 48 studies for inclusion in our meta-analysis.

Study Coding

The meta-analysis included 48 studies and 52 teacher professional learning interventions, published between 2010 and 2020.

Next, we coded all 48 studies for relevant study information. This process involved reading the descriptions of the professional learning programs being studied and making decisions about whether the descriptions suggested the programs met the Learning Forward standards.[2] To ensure appropriate coding of the standards, we had periodic check-points with Learning Forward about our coding decisions to confirm agreement. Study coding also captured information on the professional learning impacts on teacher and student outcomes. Prior to coding, a group of coders was trained by the study leads through a series of practice studies. Once trained, the coding group moved forward with the full set of studies and met on a weekly basis to discuss any questions during coding. In addition, 33% of studies were dual coded by the coders and the study leads to support ongoing consistency across coders as coding progressed.

After coding was complete, we computed "effect sizes" for each of the teacher instruction and student learning outcomes coded. Converting the study findings into effect sizes created a common metric for the study outcome

[1] We also included studies with student achievement outcomes; however, a study was still eligible if it did not report or analyze student outcomes.

[2] Given the standards revision process was ongoing at the time of coding, the wording of the 2022 standards went through some further editorial changes after coding completed. Learning Forward and AIR confirmed together that these subsequent editorial updates did not have substantive implications that would have affected the coding.

findings across the various types of outcome measures the studies used. This approach allowed us to aggregate and analyze information across the studies as a group.

Analysis

We then conducted a meta-analysis using the coded information about the professional learning programs and the computed effect sizes for teacher instruction and student learning outcomes. We examined the association between the presence of the 2011, the Areas for Deeper Exploration, and the 2022 Standards with teacher and student outcomes. We also investigated the extent to which changes in teacher instruction are associated with (or mediate) changes in student learning.

See Appendix A for further details on the study approach.

How to Read a Meta-Analysis

Key points for interpreting these results:

- **An effect size** is a standardized metric that represents the difference in an outcome between the participant and comparison groups, such as the difference in observation scores for teachers who participate in a program and those who do not. We refer to the **average effect** as the mean effect size across the programs being studied. For example, the average effect for the learning designs standard represents the mean effect size for outcomes across all programs that met the standard.
- Keep in mind that **the sizes of average effects are not directly comparable** between standards. For example, we cannot conclude that an average effect for the 2011 Resources standard (1.14) is statistically larger than that for the Outcomes standard (0.84) for teacher instruction.
- Meta-analysis can help uncover **associations, but they are *not* causal**. For example, the average effect in this study represents the relationship between a standard and outcome. However, this does not mean that the presence of the standard led to the outcome per se.

Findings

The meta-analysis included 48 studies that examined the impacts of 52 teacher professional learning programs. All studies considered program impacts on observational measures of instruction, and 34 of the studies also explored impacts on student achievement. Table B.1 in Appendix B lists the included studies and key summary details about the information each study contributed to the meta-analysis.

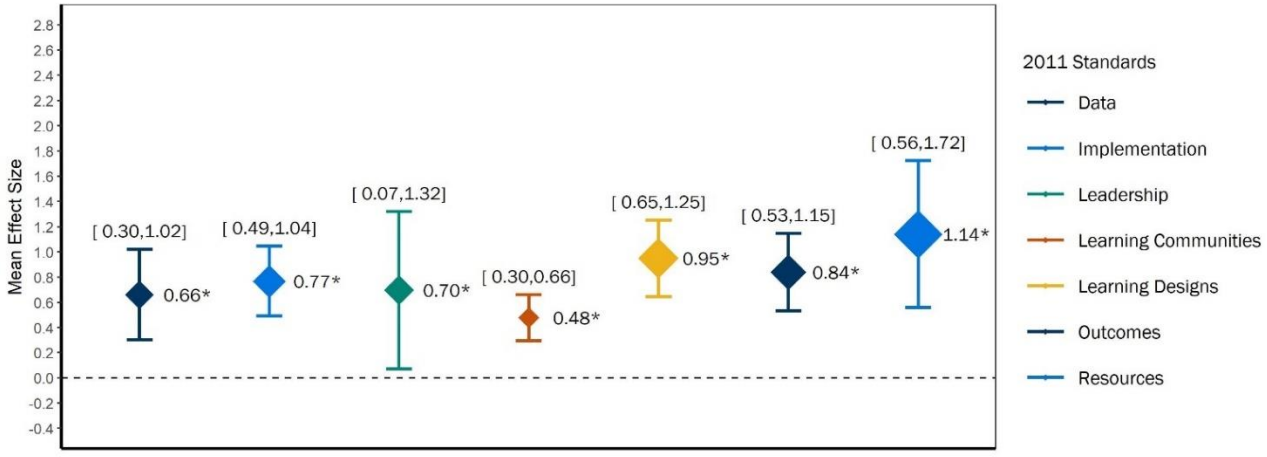
The average program effect on instruction was 0.73 standard deviations (SD), indicating a significant, positive impact of the studied professional learning programs on instruction, on average. In addition, the average program effect on student achievement was 0.09 SD, which was also statistically different from zero, indicating that the studies found on average that the professional learning was associated with improvements in student learning.

2011 Standards

To address our first research question, we considered the average program effects among programs that were identified as having features that represent the 2011 Standards. The frequency with which we found evidence of the 2011 Standards in the included studies is summarized in Table B.2 in Appendix B.

As shown in Figure 1, the findings indicated that each of the 2011 Standards were positively associated with teacher instruction. All estimates were significantly different from zero, and the average effects ranged in size from 0.48 SD (Learning Communities) to 1.14 SD (Resources). Similarly, all 2011 Standards when evident in the studies were positively and significantly associated with student achievement, with average effects ranging from 0.03 SD (Learning Communities) to 0.10 SD (Outcomes) (see Figure 2).

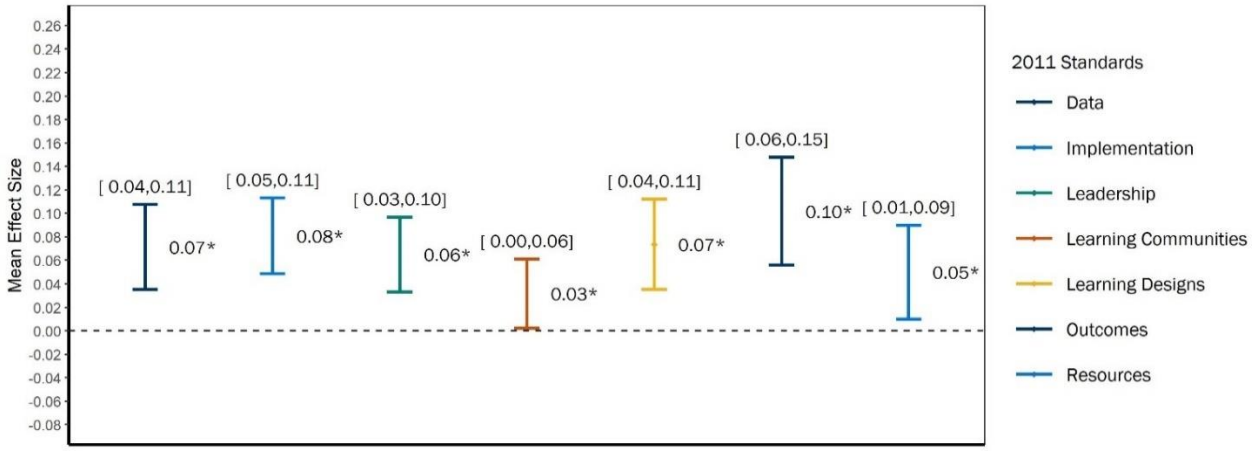
Figure 1. Average Effect Sizes of Teacher Instruction for Each 2011 Professional Learning Standard



statistically significant at α = 0.05

Notes. Mean effect sizes for the 2011 Standards based on teacher instructional observation outcomes. Forty-eight studies, 52 interventions, and 296 effect sizes were represented in the meta-regression models. The models controlled for publication type and level of random assignment.

Figure 2. Average Effect Sizes of Student Achievement for Each 2011 Professional Learning Standard



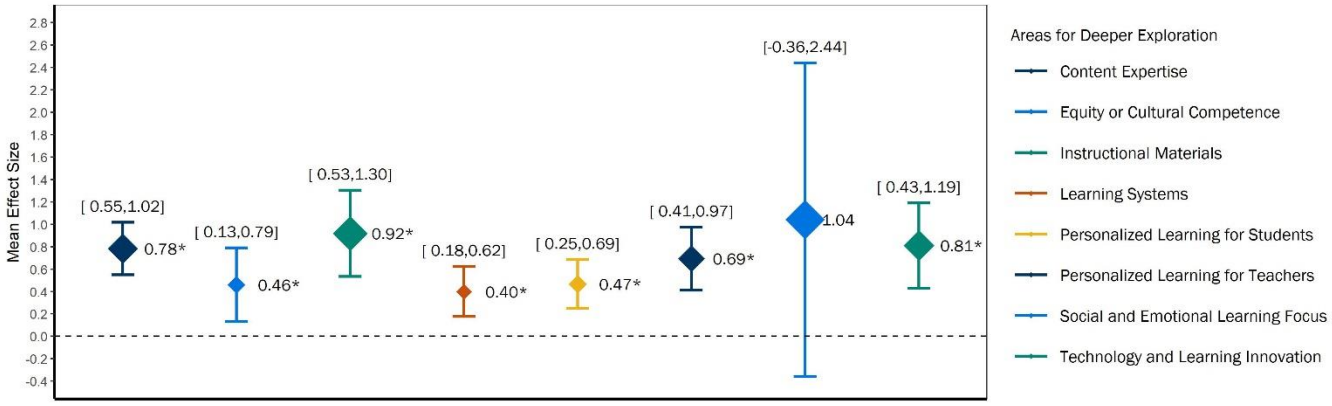
* statistically significant at α = 0.05

Notes. Mean effect sizes for the 2011 Standards based on student achievement outcomes. Thirty-four studies, 38 interventions, and 186 effect sizes were represented in the meta-regression models. The models controlled for publication type, level of random assignment, and type of achievement measure (standardized or researcher developed).

Areas for Deeper Exploration

In preparation for the 2022 revisions to the standards, Learning Forward identified eight Areas for Deeper Exploration. These areas represented topics or approaches to professional learning that were not specifically addressed in the 2011 Standards but that Learning Forward recognized as important to the field. When coded as present in studies, the Areas for Deeper Exploration also were associated with significant, positive effects on instruction and student achievement, with the exception of the Technology and Learning Innovation area, which was not associated with student achievement, and the Social and Emotional Learning (SEL) Focus area, which was not associated with either teacher instruction or student achievement (see Figures 3 and 4). We caution interpretation of analyses for the Equity, Personalized Learning for Students, and SEL Focus areas in particular, as they were represented in only a limited number of studies (see Table B.2. in Appendix B for the counts).

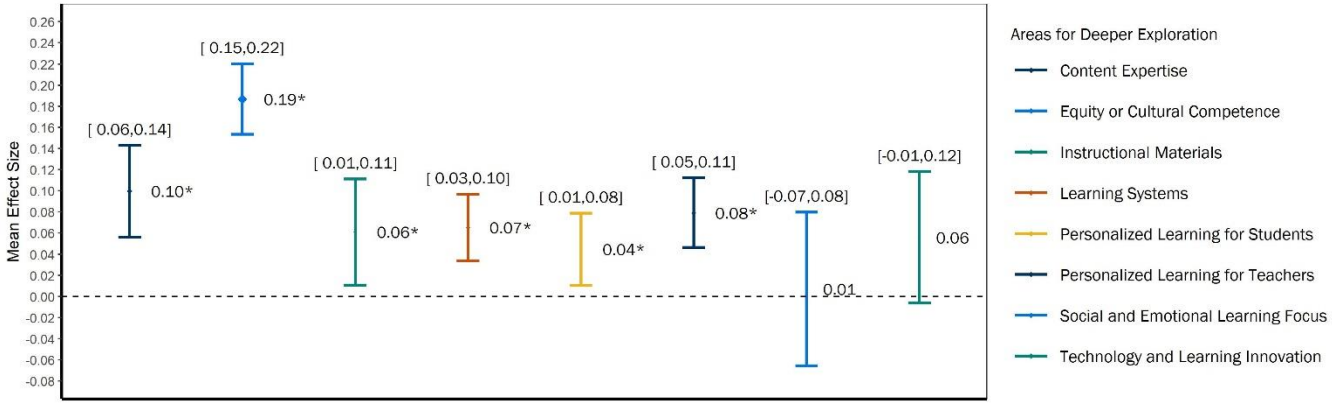
Figure 3. Average Effect Sizes of Teacher Instruction for Each Area for Deeper Exploration



* statistically significant at α = 0.05

Notes. Mean effect sizes for the Areas for Deeper Exploration based on teacher instructional observation outcomes. Forty-eight studies, 52 interventions, and 296 effect sizes were represented in the meta-regression models. The models controlled for publication type and level of random assignment.

Figure 4. Average Effect Sizes of Student Achievement for Each Area for Deeper Exploration



* statistically significant at α = 0.05

Notes. Mean effect sizes for the Areas for Deeper Exploration based on student achievement outcomes. Thirty-four studies, 38 interventions, and 186 effect sizes were represented in the meta-regression models. The models controlled for publication type, level of random assignment, and type of achievement measure (standardized or researcher developed).

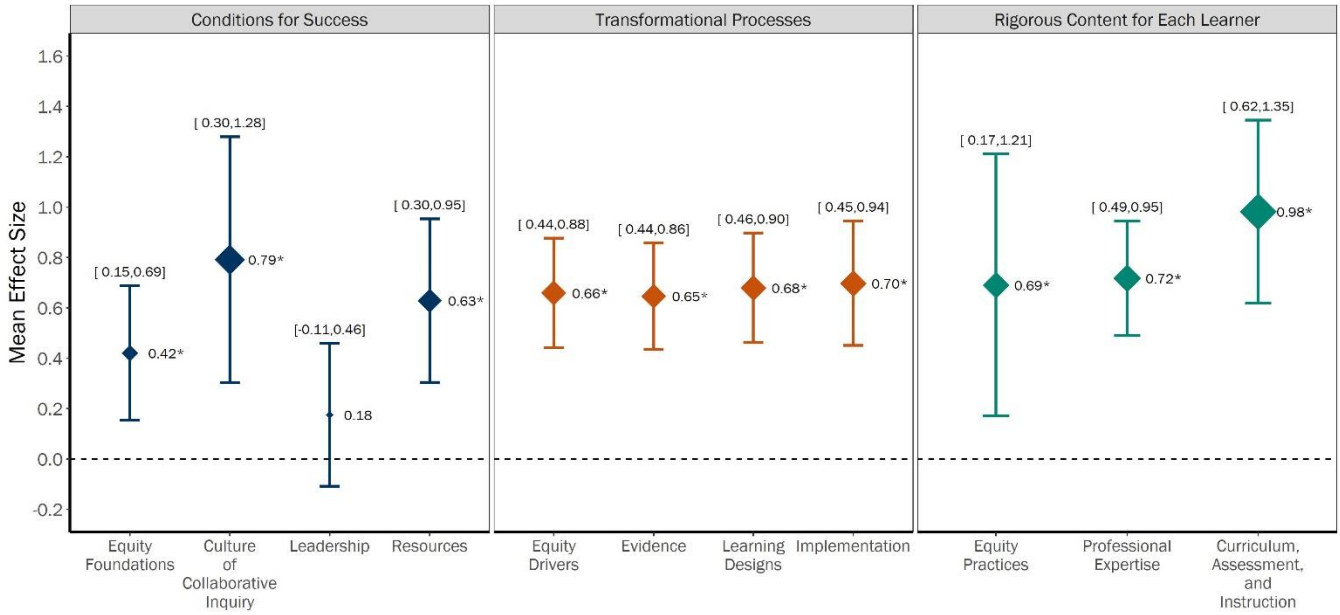
2022 Standards

We next focused on examining how features of professional learning programs consistent with the new 2022 Professional Learning Standards were associated with teacher instruction and student achievement. Table B.2 in Appendix B summarizes the number of studies that described professional learning aligned with each of the revised standards. With the reformulation of the standards, we found different levels of evidence present in the study descriptions even when standards maintained the same name in the 2022 revision. For example, because of the difference in definitions, we found nine studies with evidence of alignment to the 2011 Leadership standard but only one study aligned to the 2022 Leadership standard.

Further, studies were less often found to demonstrate evidence of the 2022 Standards under the Conditions for Success frame compared to the other frames. Only a limited number of studies provided evidence of the three equity standards or the Leadership standard. In light of the paucity of studies meeting the Leadership, Equity Drivers, and Equity Practices standards, we interpret all analyses of these standards with caution. In comparison, 28 of the 48 included studies described professional learning aligned with the Learning Designs standard, and 32 met the standard for Professional Expertise—the most commonly met of the 2022 Standards.

As shown in Figure 5, the findings indicated large, positive average effects on instruction for almost all of the 2022 Standards. All of the findings indicated positive, statistically significant average effects with the exception of a positive but non-significant effect for the revised Leadership standard. Among the statistically significant effects, the average effects on instruction ranged from 0.42 standard deviations (Equity Foundations) to 0.98 standard deviations (Curriculum, Assessment, and Instruction).

Figure 5. Average Effect Sizes of Teacher Instruction for Each 2022 Professional Learning Standard

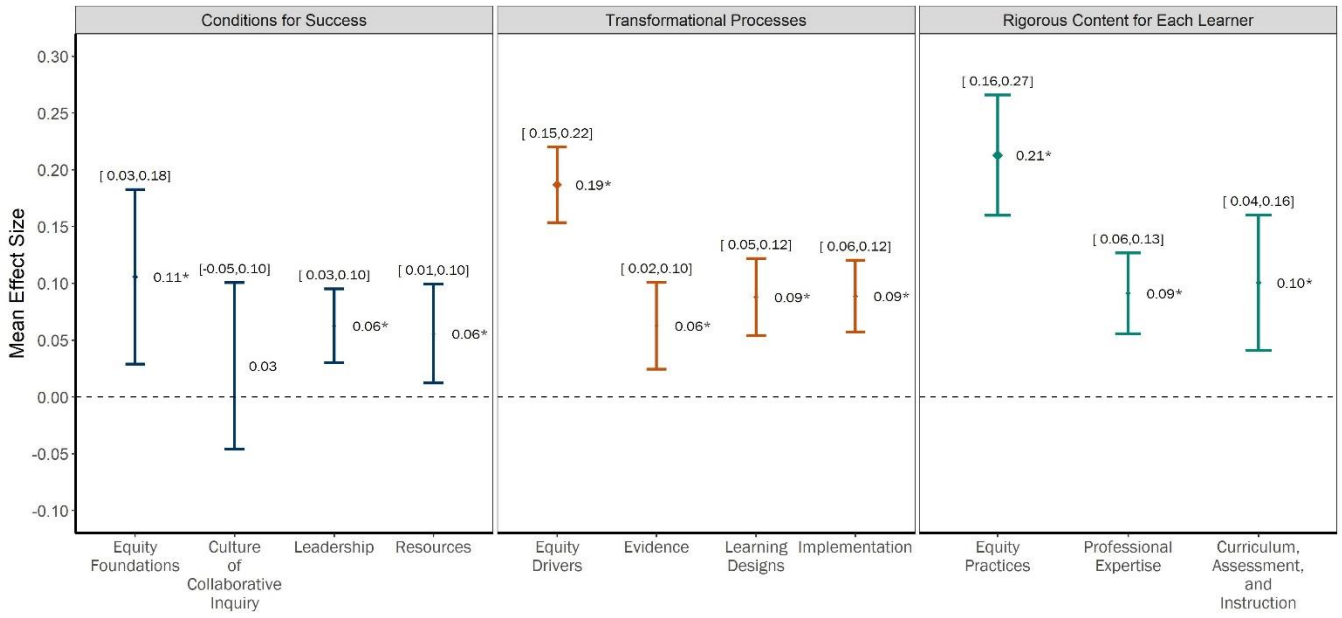


* statistically significant at α = 0.05

Notes. Mean effect sizes for the 2022 Standards based on teacher instructional observation outcomes. Forty-eight studies, 52 interventions, and 296 effect sizes were represented in the meta-regression models. The models controlled for publication type and level of random assignment.

The findings were similar when considering student achievement. Analyses (shown in Figure 6) indicated positive average effects for student achievement across all of the 2022 Standards, with statistically significant findings for all standards except the Culture of Collaborative Inquiry standard. Among the statistically significant findings, average effects ranged from 0.06 standard deviations (among each of the Leadership, Resources, and Evidence standards) to 0.21 standard deviations (Equity Practices).

Figure 6. Average Effect Sizes of Student Achievement for Each 2022 Professional Learning Standard



* statistically significant at α = 0.05

Notes. Mean effect sizes for the 2022 Standards based on student achievement outcomes. Thirty-four studies, 38 interventions, and 186 effect sizes were represented in the meta-regression models. The models controlled for publication type, level of random assignment, and type of achievement measure (standardized or researcher developed).

Gaps in the Research

To further summarize where there was available information from the included research for each of the 2022 Standards—and what the findings indicated about effects on teacher instruction and student achievement—we created "evidence gap maps" (see Figures 7 and 8). In the evidence gap maps, a circle corresponds to each standard. The number inside the circle and the circle size is proportional to the number of studies in which we found evidence of the presence of the standard. The color of each circle also corresponds to the average effect on the teacher or student outcome by each standard, with smaller effects indicated by colors closer to red and larger effects indicated by colors closer to green. For example, the Curriculum, Assessment, and Instruction standard has a mid-sized circle for teacher instruction because it was represented in 14 studies; it is also a solid green, as the standard was associated with the largest effects on teacher instruction.

Figure 7. Evidence Gap Map for the 2022 Professional Learning Standards for Teacher Instruction Outcomes

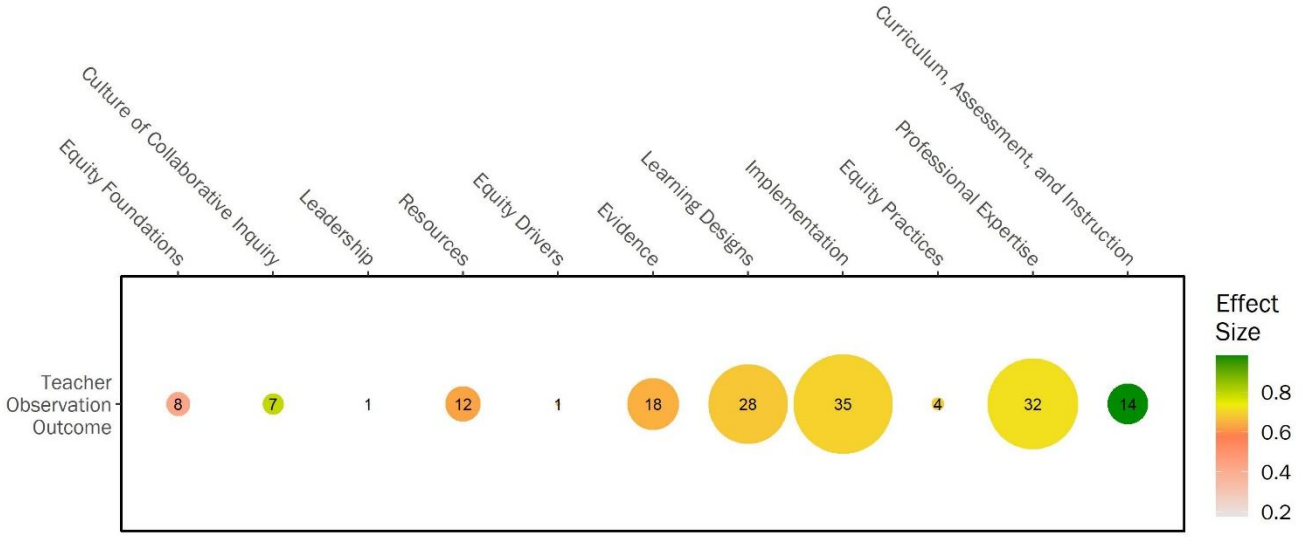
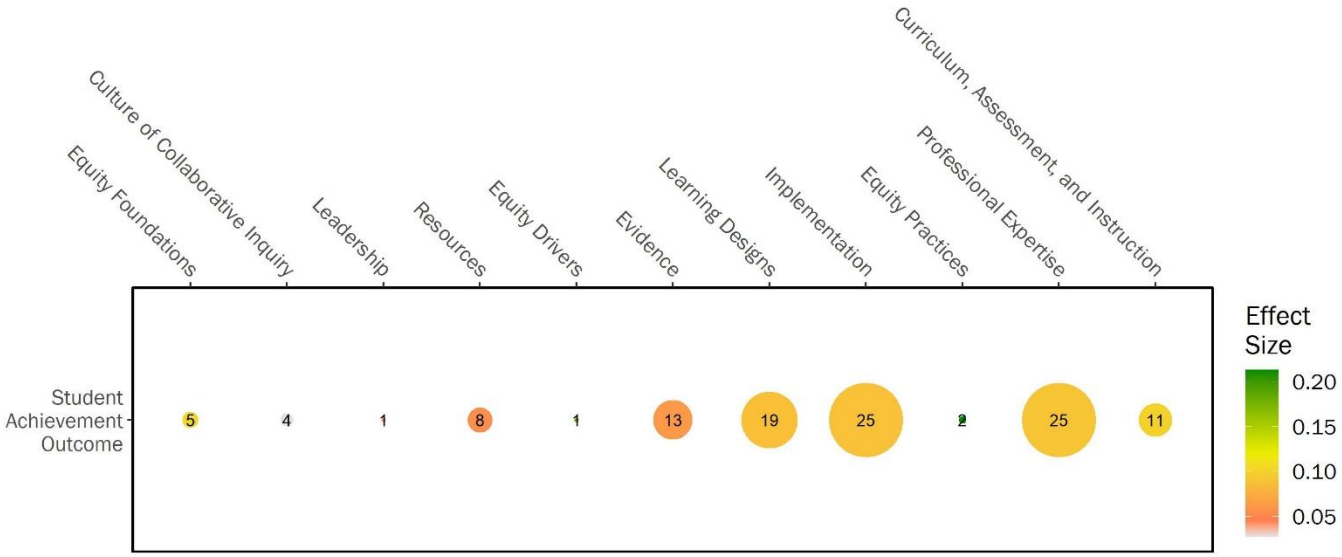


Figure 8. Evidence Gap Map for the 2022 Professional Learning Standards for Student Achievement Outcomes



The evidence gap maps bring several points into focus. First, the small circles for the standards in the Conditions for Success frame underscore how randomized field studies need to provide more information to understand the contextual factors that contribute to the impacts of professional learning. Similarly, the small circles for the equity standards visually depict the paucity of research on professional learning that explicitly addressed issues of equity, as there was little available information on the equity standards across all three frames. Third, the yellow and green colors denoting larger effect sizes suggest the particular strength of evidence of professional learning that meets the Curriculum, Assessment, and Instruction standard for supporting both teacher and student outcomes.

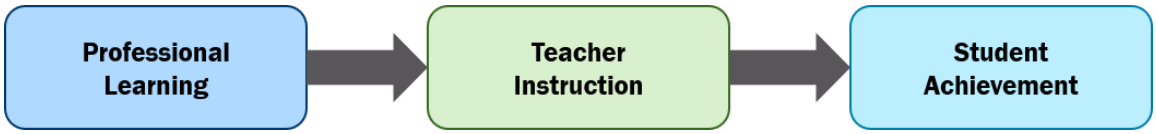
Mediation Analysis

Underlying professional learning is a theory of change that predicts professional learning will benefit both teacher instruction and student achievement outcomes, and that improvements in student achievement derive, in part, through the instructional improvements. Although this theory of change is relevant for virtually all professional learning studies, to date the field has had limited information to empirically address this theory. A meta-analysis provides an ideal opportunity to build knowledge about these theorized connections and, thus, help us understand how investments in teachers may be expected to bring about improvements for teachers and students.

To formally examine the theory of change, we conducted a mediation-meta-analysis.[3] The mediation analysis allowed us to draw on all available information across the studies and analyze all of these relationships together in a unified model. This process helped us to understand, in particular, the effects that teacher professional learning has on student achievement and the portion of the effects on student achievement that can be attributed to observed changes in teacher instruction.

The analyses revealed two key points. First, as shown in Figure 9, when we examined how these relationships work together as a whole, we found empirical evidence to support the theory of change that improvements in instruction were positively and significantly associated with improvements in student achievement. Second, once the analyses accounted for the instructional improvements associated with professional learning, the improvements in student achievement were explained through the improvements in teacher instruction that came from the professional learning.

Figure 9. Effects of Teacher Professional Learning on Student Achievement as Mediated by Changes in Teacher Instruction



[3] See Appendix A. Mediation Models, for further details.

Discussion

This study meta-analyzed information from 48 studies that used a rigorous approach to testing teacher professional learning programs. The findings consistently indicated that evidence of program alignment with the Learning Forward professional learning standards was associated with improved teacher instruction and student achievement. The study also shed light on where rigorous studies of teacher professional learning bring little information to bear and, thus, suggest opportunities for further research.

While this meta-analysis helps contextualize the Learning Forward professional learning standards within studies of randomized experiments of teacher professional learning, we emphasize that the standards are also based on factors beyond this specific type of research evidence. Some standards reflect expert input and values, particularly those standards that are relatively newer considerations in the field of professional learning. For example, equity in education has received heightened attention in recent years around its urgency. Learning Forward aimed to reflect the importance of equity in the professional learning standards despite relatively few research studies with that focus. Second, the standards have foundation upon other forms of research beyond randomized field studies. For example, descriptive and qualitative studies have long identified that importance of district and school leadership for creating a system that enables effective teacher professional learning (Drago-Severson, 2012; Whitworth & Chiu, 2015). Accordingly, this meta-analysis is one of multiple sources of research and knowledge from the field that has contributed to developing the revised professional learning standards.

How These Findings Compare to—and Build on—Prior Research

Although this is the first meta-analysis study to examine the Learning Forward standards, we can contextualize our general findings in reference to earlier meta-analyses focused on teacher professional learning. Considering effects on teacher instruction, our average effects are higher at 0.74 SD than the average effect of 0.42 from an earlier meta-analysis of professional learning that included research as far back as 1995 (Garrett et al., 2019) and average effect of 0.49 from a prior meta-analysis of coaching programs that mostly targeted literacy (Kraft et al., 2018). Given the variation in effects on instruction found in all three meta-analyses, we would caution against interpreting that the average effects on instruction are definitively different across the studies. We would instead encourage interpreting the cumulative findings as further evidence that when teachers engage in professional learning, they are able to improve their instruction in the ways they are asked to do so.

The mean effects of 0.09 SD on student achievement are more modest than the average achievement effect of 0.18 in the Kraft et al. coaching study, as well as the average achievement effect of 0.21 SD in a meta-analysis of STEM-focused curriculum and professional development studies (Lynch et al., 2019). Although the average effects on student achievement found in this study may seem modest, prior research supports that they are substantively meaningful. Kraft (2020) analyzed information on effects of educational interventions among 747 research studies and concluded that among randomized field trials using standardized achievement outcomes, an effect as low at 0.05 SD should be considered medium-sized, given the difficulty in raising student achievement through any approach. By Kraft's benchmarks, the findings from this study based on rigorous research demonstrate that investments in professional learning can yield meaningful improvements in student achievement.

The findings from this study based on rigorous research demonstrate that investments in professional learning can yield meaningful improvements in student achievement.

Further, the criteria used to identify studies relevant to include in the analysis may have limited our ability to fully capture the effects of professional learning on student achievement. Given the focus of this work on professional learning standards for teachers, we required that studies include analyses of impacts on teacher instruction. Because of this requirement, an additional set of studies—randomized field trials that published impacts on student achievement but not on teacher instruction—are not represented in our sample. We note that prior meta-analyses on teacher professional learning that did not require an analysis of impact on instruction have found larger, positive mean effects for student achievement (Kraft et al., 2018; Lynch et al., 2019). As such, our findings related to student achievement may present a conservative, more minimal estimate of potential benefits to students. Similarly, it is possible that requiring studies to include an analysis of impact on teacher instruction may have contributed to the larger effects on instruction found in this study compared to others, particularly if it is less likely to find null or negative impacts on instruction in the published research. However the larger mean effects for instruction compared to the Garrett et al. study (2019), which used the same inclusion criteria, would suggest this is not the case.

Limitations to Acknowledge

It is important to acknowledge the inherent limitations to what a meta-analysis can inform. For the purposes of this type of analysis, we focused on a specific type of research (randomized field trials), which means the analyses do not draw on the rich findings from other types of important research on professional learning. Also, this study includes 48 studies, making it a robust meta-analysis; however, there are still places where the amount of available data is limited. In addition, a meta-analysis does not allow researchers to disentangle the contributions of specific program features in programs that contain multiple components. For example, a study may evaluate the impact of a program that includes both individualized coaching and a focus on using data to inform instruction, but researchers cannot distinguish the unique contribution of these two factors to instructional or student learning outcomes. Last, it is important to recognize that the Learning Forward standards are relevant for all educators and not just classroom teachers. This study focuses only on research relevant to classroom teachers, and further investigation into the relationship of professional learning that embodies the standards with outcomes for other groups of educators is warranted.

Future Directions for Research

The broader context of teacher professional learning program implementation | This study has highlighted the need for additional research in several key areas of teacher professional learning. One such area is increased understanding of the broader systems and contexts in which professional learning is tested. This need was evidenced by the lack of information we were able to gather from the studies that was relevant to the Conditions for Success frame in the 2022 Standards. In part, this lack of information about the broader implementation context may reflect a limited emphasis on attending to contextual information in publications of randomized field trials—despite the importance of understanding context for understanding professional learning effectiveness (Hill et al., 2013; Hill et al., 2016). However the limited descriptions relevant to the Conditions for Success frame also surfaces how little the professional learning programs studied directly incorporated intentional connections within their implementation settings. Practices such as helping educators build a culture of shared responsibility for student learning or supporting leaders to simultaneously enable both individual and collective capacity building were rarely part of the professional learning being studied. This suggests an opportunity for further investigation.

Teacher professional learning that addresses equity | Similarly, we found little evidence of professional learning aligned with the equity standards. As the COVID-19 pandemic has shone a brighter light than ever on areas in

which equity is lacking (Donohue & Miller, 2020; National Academies of Sciences, Engineering, and Medicine, 2020), the field has a renewed urgency for addressing equity through teacher professional learning. It is notable that when we found evidence of professional learning embodying either the equity standards or the standards within the Conditions for Success frame, the findings indicated promise for teacher and student outcomes. To build on these signs of promise, researchers and program developers together have an opportunity to more deeply incorporate both equity and features of the broader context into professional learning, and study them within rigorous research.

Scaling, cost, and other aspects of teacher professional learning programs | Finally, we recognize the need for increased understanding of other topics germane to teacher professional learning—or any educational intervention. For example, there is growing interest within the research field in understanding the professional learning experiences of the "comparison group" in randomized field studies, the costs of the professional learning being studied, and data to inform the potential for scaling a given program. Although our research questions did not address these areas, we still noted that we would have limited ability to learn about them from the included studies. It will be important for research to provide information on these topics as we look ahead.

Studies Included in the Meta-Analysis

Babinski, L. M., Amendum, S. J., Knotek, S. E., Sánchez, M., & Malone, P. (2018). Improving young English learners' language and literacy skills through teacher professional development: A randomized controlled trial. *American Educational Research Journal*, *55*(1), 117–143. doi: 10.3102/0002831217732335

Baker, D. L., Santoro, L., Biancarrosa, G., & Baker, S. (2015). *Effects of quality of instruction on student vocabulary and comprehension during read alouds.* SREE Spring 2015 Conference Abstract, Washington, DC.

Blazar, D., & Kraft, M. A. (2015). Exploring mechanisms of effective teacher coaching: A tale of two cohorts from a randomized experiment. *Educational Evaluation and Policy Analysis*, *37*(4), 542–566. doi: 10.3102/0162373715579487

Bos, J. M., Sanchez, R. C., Tseng, F., Rayyes, N., Ortiz, L., & Sinicrope, C. (2012). E*valuation of Quality Teaching for English Learners (QTEL) professional development* (NCEE 2012-4005). National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Brown, J. L., Jones, S. M., LaRusso, M. D., & Aber, L. (2010). Improving classroom quality: Teacher influences and experimental impacts of the 4Rs program. *Journal of Educational Psychology*, *102*(1), 153–167.

Bruns, B., Costa, L., & Cunha, N. (2017). Through the looking glass: can classroom observation and coaching improve teacher performance in Brazil? *Economics of Education Review, 64*, 214–250.

Connor, C. M., Morrison, F. J., Schatschneider, C., Toste, J. R., Lundblom, E., Crowe, E. C., & Fishman, B. (2011). Effective classroom instruction: Implications of child characteristics by reading instruction interactions on first graders' word reading achievement. *Journal of Research on Educational Effectiveness, 4*(3), 173–207. doi: 10.1080/19345747.2010.510179

Cordray, D., Pion, G., Brandt, C., Molefe, A., & Toby, M. (2012). *The impact of the Measures of Academic Progress (MAP) Program on student reading achievement: Final report* (NCEE 2013-4000). National Center for Education Evaluation and Regional Assistance.

DeCesare, D., McClelland, A., & Randel, B. (2017). *Impacts of the retired mentors for new teachers program.* U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Central.

Doabler, C. T., Nelson, N. J., Kosty, D. B., Fien, H., Baker, S. K., Smolkowski, K., & Clarke, B. (2014). Examining teachers' use of evidence-based practices during core mathematics instruction. *Assessment for Effective Intervention*, *39*(2), 99–111.

Dolfin, S., Richman, S., Choi, J., Streke, A., DeSaw, C., Demers, A., & Poznyak, D. (2019). *Evaluation of the Teacher Potential Project.* Mathematica.

Fabiano, G. A., Reddy, L. A., & Dudek, C. M. (2018). Teacher coaching supported by formative assessment for improving classroom practices. *School Psychology Quarterly*, *33*(2), 293–304. http://dx.doi.org/10.1037/spq0000223

Faraclas, K. L. (2018). A professional development training model for improving co-teaching performance. *International Journal of Special Education*, *33*(3), 524–540.

Garet, M. S., Wayne, A. J., Stancavage, F., Taylor, J., Eaton, M., Walters, K., Song, M., Brown, S., Hurlburt, S., Zhu, P., Sepanik, S., Doolittle, F., & Warner, E. (2011). *Middle school mathematics professional development impact study: Findings after the second year of implementation* (NCEE 2011-4024). National Center for Education Evaluation and Regional Assistance.

Garet, M. S., Wayne, A. J., Stancavage, F., Taylor, J., Walters, K., Song, M., Brown, S., Hurlburt, S., Zhu, P., Sepanik, S., Doolittle, F., & Warner, E. (2010). *Middle school mathematics professional development impact study: Findings after the first year of implementation* (NCEE 2011-4024). National Center for Education Evaluation and Regional Assistance.

Garet, M. S., Heppen, J. B., Walters, K., Parkinson, J., Smith, T. M., Song, M., Garrett, R., Yang, R., & Borman, G. D. (2016). *Focusing on mathematical knowledge: The impact of content-intensive teacher professional development* (NCEE 2016-4010). National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Garet, M. S., Wayne, A. J., Brown, S., Rickles, J., Song, M., & Manzeske, D. (2017). *The impact of providing performance feedback to teachers and principals.* National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Gersten, R., Dimino, J., Jayanthi, M., Kim, J. S., & Santoro, L. E. (2010). Teacher study group: Impact of the professional development model on reading instruction and student outcomes in first grade classrooms. *American Educational Research Journal*, *47*(3), 694–739. doi: 10.3102/0002831209361208

Goldman, S. R., Greenleaf, C., Yukhymenko-Lescroart, M., Brown, W., Ko, M., Emig, J., George, M., Wallace, P., Blum, D., & Britt, M. A. (2019). Explanatory modeling in science through text-based investigation: Testing the efficacy of the Project READI intervention approach. *American Educational Research Journal*, *56*(4), 1148–1216.

Goodson, B., Wolf, A., Bell, S., Turner, H., & Finney, P. B. (2010). *The effectiveness of a program to accelerate vocabulary development in kindergarten (VOCAB)* (NCEE 2010-4014). National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Granger, E. M., Bevis, T. H., Saka, Y., Southerland, S. A., Sampson, V., & Tate, R. L. (2012). The efficacy of student-centered instruction in supporting science learning. *Science*, *338*(105), 105–108. doi: 10.1126/science.1223709

Granger, E. M., Bevis, T. H., Southerland, S. A., Saka, Y., & Ke, F. (2019). Examining features of how professional development and enactment of educative curricula influences elementary science teacher learning. *Journal of Research in Science Teaching*, *56*, 348–370. doi: 10.1002/tea.21480

Grigg, J., Kelly, K. A., Gamoran, A., & Borman, G. D. (2013). Effects of two scientific inquiry professional development interventions on teaching practice. *Educational Evaluation and Policy Analysis*, *35*(1), 215–229.

Jacob, R., Hill, H., & Corey, D. (2017). The impact of a professional development program on teachers' mathematical knowledge for teaching, instruction, and student achievement. *Journal of Research on Educational Effectiveness*, *10*(2), 379–407.

Jitendra, A. K., Harwell, M. R., Dupuis, D. N., Karl, S. R., Lein, A. E., Simonson, G., & Slater, S. C. (2015). Effects of a research-based mathematics intervention to improve seventh-grade students' proportional problem solving: A cluster randomized trial. *Journal of Educational Psychology*, *107*, 1019–1034. doi: 10.1037/edu0000039

Jitendra, A. K., Harwell, M. R., Karl, S. R., Slater, S. C., Simonson, G. R., & Nelson, G. (2016). *A replication study to evaluate the effects of schema-based instruction on middle school students' proportional problem-solving performance.* SREE Spring 2016 Conference Abstract, Washington, DC.

Johnson, A., Galloway, C., Friedlander, E., & Goldenberg, C. (2019). Advancing educational quality in Rwanda: Improving teachers' literacy pedagogy and print environments. *International Journal of Educational Research, 98*, 134–145.

Johnson, C. C., & Fargo, J. D. (2010). Urban school reform enabled by transformative professional development: Impact on teacher change and student learning of science. *Urban Education*, *45*(1), 4–29.

Jukes, M. C., Turner, E. L., Dubeck, M. M., Halliday, K. E., Inyega, H. N., Wolf, S., Zuilkowski, S. S., & Brooker, S. J. (2017). Improving literacy instruction in Kenya through teacher professional development and text messages support: A cluster randomized trial. *Journal of Research on Educational Effectiveness*, *10*(3), 449–481.

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

Kim, J. S., Olson, C. B., Scarcella, R., Kramer, J., Pearson, M., van Dyk, D., Collins, P., & Land, R. (2011). A randomized experiment of a cognitive strategies approach to text-based analytical writing for mainstreamed Latino English language learners in grades 6 to 12. *Journal of Research on Educational Effectiveness*, *4*(3), 231–263.

Kraft, M., & Blazar, D. (2016). Individualized coaching to improve teacher practice across grades and subjects: New experimental evidence. *Educational Policy, 31*. doi: 10.1177/0895904816631099

Lara-Alecio, R., Irby, B., Tong, F., Guerrero, C., & Cajiao-Wingenbach, L. (2018). *English Language and Literacy Acquisition-Validation (ELLA-V) i3 Evaluation (Valid 22) final report.* Center for Research and Reform in Education.

LaVenia, M. (2016). *Mathematics formative assessment system: Testing the theory of action based on the results of a randomized field trial* [Doctoral dissertation; Florida State University]. http://purl.flvc.org/fsu/fd/FSU_2016SU_LaVenia_fsu_0071E_13439

Matsumura, L. C., Garnier, H. E., & Spybrook, J. (2013). Literacy coaching to improve student reading achievement: A multi-level mediation model. *Learning and Instruction*, *25*, 35–48. http://dx.doi.org/10.1016/j.learninstruc.2012.11.001

Meyers, C. V., Molefe, A., Brandt, W. C., Zhu, B., & Dhillion, S. (2016). Impact results of the eMINTS professional development validation study. *Educational Evaluation and Policy Analysis*, *38*(3), 455–476. doi: 10.3102/0162373716638446

Motoca, L. M., Farmer, T. W., Hamm, J. V., Byun, S. Y., Lee, D. L., Brooks, D. S., Rucker, N., & Moohr, M. M. (2014). Directed consultation, the SEALS Model, and teachers' classroom management. *Journal of Emotional and Behavioral Disorders*, *22*(2), 119–129.

Murray, D. W., Rabiner, D. L., Kuhn, L., Pan, Y., & Sabet, R. F. (2018). Investigating teacher and student effects of the Incredible Years Classroom Management Program in early elementary school. *Journal of School Psychology*, *67*, 119–133.

Nelson-Walker, N. J., Fien, H., Kosty, D. B., Smolkowski, K., Smith, J. L. M., & Baker, S. K. (2013). Evaluating the effects of a systemic intervention on first-grade teachers' explicit reading instruction. *Learning Disability Quarterly*, *36*(4), 215–230.

Nugent, G., Kunz, G., Houston, J., Kalutskaya, I., Wu, C., Pedersen, J., Lee, S., DeChenne, S. E., Luo, L., & Berry, B. (2016). *The effectiveness of technology-delivered science instructional coaching in middle and high school* (Working Paper). National Center for Research on Rural Education.

Ottmar, E. R., Rimm-Kaufman, S. E., Berry, R. Q., & Larsen, R. A. (2013). Does the responsive classroom approach affect the use of standards-based mathematics teaching practices? Results from a randomized controlled trial. *The Elementary School Journal*, *113*(3), 434–457.

Ottmar, E. R., Rimm-Kaufman, S. E., Larsen, R. A., & Berry, R. Q. (2015). Mathematical knowledge for teaching, standards-based mathematics teaching practices, and student achievement in the context of the "Responsive Classroom Approach." *American Educational Research Journal*, *52*(4), 787–821. doi: 10.3102/0002831215579484

Parkinson, J., Salinger, T., Meakin, J., & Smith, D. (2015). *Results from a three-year i3 impact evaluation of the Children's Literacy Initiative (CLI): Implementation and impact findings of an intensive professional development and coaching program.* American Institutes for Research.

Penuel, W. R., Gallagher, L. P., & Moorthy, S. (2011). Preparing teachers to design sequences of instruction in Earth systems science: A comparison of three professional development programs. *American Educational Research Journal*, *48*(4), 996–1025. https://doi.org/10.3102/0002831211410864

Reddy, L. A., Dudek, C. M., & Lekwa, A. (2017). Classroom strategies coaching model: Integration of formative assessment and instructional coaching. *Theory Into Practice*, *56*, 46–55. doi: 10.1080/00405841.2016.1241944

Reinke, W. M., Herman, K. C., & Dong, R. A. (2018). The Incredible Year teacher classroom management program: Outcomes from a group randomized trial. *Prevention Science*, *19*, 1043–1054. https://doi.org/10.1007/s11121-018-0932-3

Rimm-Kaufman, S. E., Larsen, R. A. A., Baroody, A. E., Curby, T. W., Ko, M., Thomas, J. B., Merritt, E. G., Abry, T., & DeCoster, J. (2014). Efficacy of the responsive classroom approach: Results from a 3-year, longitudinal randomized controlled trial. *American Educational Research Journal*, *51*(3), 1–37. doi: 10.3102/0002831214523821

Sailors, M., & Price, L. (2010). Professional development that supports the teaching of cognitive reading strategy instruction. *The Elementary School Journal*, *110*, 301–322.

Santagata, R., Kersting, N., Givvin, K. B., & Stigler, J. W. (2011). Problem implementation as a lever for change: An experimental study of the effects of a professional development program on students' mathematics learning. *Journal of Research on Educational Effectiveness*, *4*(1), 1–24. doi: 10.1080/19345747.2010.498562

Simmons, K. R. (2010). Effects of a rubric-driven assessment on teacher performance with regard to behavioral management skills. (Unpublished doctoral dissertation). St John's University School of Education.

Supovitz, J. A. (2013). *The linking study: An experiment to strengthen teachers' engagement with data on teaching and learning* (CPRE Working Papers). Consortium for Policy Research in Education. http://repository.upenn.edu/cpre_workingpapers/3

Supovitz, J., & Sirinides, P. (2017). The linking study: An experiment to strengthen teachers' engagement with data on teaching and learning. *American Journal of Education*, *124*(2), 161–189.

Tang, S., Tong, F., Irby, B. J., Lara-Alecio, R., & Guerrero, C. (2020). Fidelity of implementation in a randomized controlled trial study: The effect of virtual professional development on bilingual teachers. *Bilingual Research Journal*, *43*(1), 111–124. doi: 10.1080/15235882.2019.1711268

Taylor, J. A., Getty, S. R., Kowalski, S. M., Wilson, C. D., Carlson, J., & Van Scotter, P. (2015). An efficacy trial of research-based curriculum materials with curriculum-based professional development. *American Educational Research Journal, 52*(5), 984–1017.

Taylor, J. A., Kowalski, S. M., Getty, S. R., Wilson, C. D., & Carlson, J. (2013). *The effects of research-based curriculum materials and curriculum-based professional development on high school science achievement: Results of a cluster-randomized trial.* SREE Spring 2013 Conference Abstract, Washington, DC. https://eric.ed.gov/?id=ED564104

Taylor, J. A., Kowalski, S. M, Getty, S. R., Wilson, C. D., & Carlson, J. (2011). *The impact of curriculum-based professional development on science instruction: Results from a cluster-randomized trial.* SREE Fall 2011 Conference Abstract, Washington, DC. https://eric.ed.gov/?id=ED528838

Tong, F., Irby, B. J., Lara-Alecio, R., Guerrero, C., Tang, S., & Sutton-Jones, K. L. (2019). The impact of professional learning on in-service teachers' pedagogical delivery of literacy-infused science with middle school English learners: A randomised controlled trial study in the U.S. *Educational Studies*, *45*(5), 533–553. doi: 10.1080/03055698.2018.1509776

Vadasy, P. F., Sanders, E. A., & Logan Herrera, B. (2015). Efficacy of rich vocabulary instruction in fourth-and fifth-grade classrooms. *Journal of Research on Educational Effectiveness*, *8*(3), 325–365.

Works Cited

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis* (1st ed.). Wiley.

Cheung, M. W. L. (2014). metaSEM (Version 0.9.4). [Software]. http://courses.nus.edu.sg/course/psycwlm/Internet/metaSEM/

Dolfin, S., Richman, S., Choi, J., Streke, A., DeSaw, C., Demers, A., & Poznyak, D. (2019). *Evaluation of the Teacher Potential Project.* Mathematica.

Donohue, J. M., & Miller, E. (2020). COVID-19 and school closures. *JAMA, 324*(9), 845–847. https://doi.org/10.1001/jama.2020.13092

Drago-Severson, E. (2012). New opportunities for principal leadership: Shaping school climates for enhanced teacher development. Teachers College Records, 114(3), 1–44.

Garrett, R., Citkowicz, M., & Williams, R. (2019). How responsive is a teacher's classroom practice to intervention? A meta-analysis of randomized field studies. *Review of Research in Education, 43*(1), 106–137. https://doi.org/10.3102/0091732X19830634

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational and Behavioral Statistics*, *6*, 107–128.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* Academic Press.

Hedges, L. V. (2007). Effect sizes in cluster-randomized designs. *Journal of Educational and Behavioral Statistics*, *32*(4), 341–370. doi: 10.3102/1076998606298043

Hedges, L. V. (2011). Effect sizes in three-level cluster-randomized experiments. *Journal of Educational and Behavioral Statistics, 36*(3), 346–380. doi: 10.3102/1076998610376617

S. A. (2019). *Revising the Learning Forward Standards for Professional Learning.* Dallas, TX: WestEd and Learning Forward.

Hill, H. C., Beisiegel, M., & Jacob, R. (2013). Professional development research: Consensus, crossroads and challenges. *Educational Researcher, 42*(9), 476–487.

Hill, H. C., Corey, D., & Jacob, R. (2016). *Dividing by zero: Exploring null results in a mathematics professional development program.* National Science Foundation.

Kraft, M. (2020). Interpreting effect sizes of educational interventions. *Educational Researcher, 49*(4), 241–253.

Kraft, M. A., Blazar, D., & Hogan, D. (2018). The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence. *Review of Educational Research*, *88*(4), 547–588.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis.* SAGE.

Lynch, K., Hill, H. C., Gonzalez, K., & Pollard, C. (2019). Strengthening the research base that informs STEM instructional improvement efforts: A meta-analysis. *Educational Evaluation and Policy Analysis, 41*(3), 260–293. doi: 10.3102/0162373719849044

National Academies of Sciences, Engineering, and Medicine. (2020). *Reopening K-12 schools during the COVID-19 pandemic: Prioritizing health, equity, and communities*. The National Academies Press. https://doi.org/10.17226/25858

Pustejovsky, J. (2018). clubSandwich: Cluster-robust (sandwich) variance estimators with small-sample corrections (R package version 0.3.2). https://CRAN.R-project.org/package=clubSandwich

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48.

Whitworth, B.A. & Chiu, J.L. (2015). Professional development and teacher change: The missing leadership link. *Journal of Science Teacher Education, 26*(2), 121–137. doi: 10.1007/s10972-014-9411-2

Wilson, S. J., Polanin, J. R., & Lipsey, M. W. (2016). Fitting meta-analytic structural equation models with complex datasets. *Research Synthesis Methods, 7*, 121–139. doi: 10.1002/jrsm.1199

Appendix A. Methods

This appendix outlines the methods used to conduct a systematic review and meta-analysis on professional development (PD) programs and how program elements are associated with teacher instruction and student learning outcomes. It contains (1) a definition of the eligibility criteria, (2) an overview of the systematic review process, and (3) a description of the meta-analytic approach and how to interpret the resulting estimates.

Eligibility Criteria

To be eligible for inclusion in this systematic review and meta-analysis, primary studies needed to meet the following criteria:

1. The study was published between January 2010 and February 2020.
2. The study focused on any grade kindergarten through Grade 12.
3. The study tested a teacher professional learning program directed at improving instruction.
4. The study used a random-assignment research design, with randomization taking place at the teacher level or higher.
5. The study included a business-as-usual control group.
6. The study included an impact analysis on an observation-based measure of teacher instruction. If available, the study would also include measures of student achievement, such as assessments (standardized or researcher-developed), although this was not part of the inclusion criteria.
7. The study was written in English, although the study could be conducted outside of the United States.
8. The study provided sufficient information to calculate an effect size estimate and variance for each outcome to be meta-analyzed.

Systematic Review

Our literature search and retrieval process for our systematic review of PD studies is presented in Table B1. We describe the process in detail in the following subsections.

Literature Search

We first conducted searches in multiple electronic research databases. The search was limited to English language–only studies, studies including kindergarten through Grade 12, and studies published between 2010 and 2020. We searched for keywords to identify studies focused on teacher professional learning using a randomized field trial approach. We then removed duplicate studies across databases and confirmed that studies used a randomization design, which yielded 1,350 studies.

Dual Screening

We then conducted a screening process to confirm which studies met our criteria. Each study was screened by two researchers, who reviewed the titles, then the study abstracts, and finally the full texts to confirm that each study met the eligibility criteria. All screening took place in a Microsoft Excel database created specifically for the project. Screeners answered questions developed by the authors using the defined eligibility criteria. A “No”

response from reviewers to any of the questions resulted in the study being excluded from further review. If reviewers responded with either "Yes" or "Do Not Know" to all questions at a given stage, the study moved along to the next stage. Studies that made it through all three stage of screening moved into the coding phase.

Coding

To examine teacher and student outcomes, we first coded each study for information related to the intervention, sample, and outcomes studied. Before coding began, the research team met with Learning Forward to determine what intervention characteristics to code, specifically the codes related to the 2011 and 2022 Standards and areas for deeper exploration. The remaining codes focused on extracting core descriptive information and characteristics that could explain, at least in part, observed heterogeneity in the effects of professional learning (i.e., potential moderators). This information included study-level data related to publication status; sample characteristics, such as sample size and demographics; outcome measure information, such as type and domain; and summary statistics to calculate effect sizes.

A total of eight coders coded the 48 studies included in the meta-analysis. Four of the coders served as the primary coders who coded all the studies from start to finish (i.e., from study information to effect size information). The four remaining coders served as dual coders to examine coder drift: two coders dual coded 33% of the studies from start to finish, and two coders dual coded 100% of the intervention codes. Any coding discrepancies were discussed by the four coders (primary, secondary, and the two dual intervention coders) until a final decision was made regarding the appropriate code. The dual coders (also the meta-analysis leads and trainers) met with the coders on a weekly basis to discuss any questions and prevent coding drift over time.

All coding was completed in an Access database developed for this meta-analysis. The Access database had a hierarchical structure such that study-level information was coded first, followed by the intervention name and corresponding characteristics, then sample and setting information, outcome information, and finally effect size information.

Meta-Analysis

Computing Effect Sizes

We computed effect sizes to provide a common metric for synthesis across studies that measure outcomes on different scales. Effect sizes encode both the direction and the magnitude of the relationship between intervention and outcomes (Hedges & Olkin, 1985; Lipsey & Wilson, 2001). Specifically, we computed the standardized mean difference (SMD) effect size for all teacher and student outcomes reported in each study. We used the summary statistics reported in the studies to compute the SMDs (e.g., means and standard deviations, *t* tests, *F* tests, χ^2 tests, regression model estimates, and effect sizes in other metrics). The equations for calculating the SMD, or converting other effect size metrics to the SMD, can be found in Borenstein et al., (2009).

We applied two adjustments to the SMDs and their variances. First, we used Hedges's (1981) small sample bias correction to the effect size estimate to the account for small studies. Second, we adjusted the effect size variances for clustering when the level of random assignment was at the cluster level (e.g., teachers or schools were randomly assigned to conditions), using formulas provided by Hedges (2007, 2011).

Meta-Analytic Models

We used mixed-effects meta-regression models to examine teacher and student outcomes and to investigate sources of effect heterogeneity, or how effects differ based on various intervention, sample, and outcome characteristics (Borenstein et al., 2009). For each model (defined in the following subsections), we controlled for a set of methods moderators. For teacher outcomes, we controlled for publication status (whether a study was published in a journal) and level of random assignment (teacher or school level). For student outcomes, we controlled for publication status, level of random assignment, and outcome measure type (standardized or researcher developed).

All models were estimated using restricted maximum likelihood with the *metafor* package in the statistical software R (Viechtbauer, 2010). To account for effect size dependencies (i.e., multiple effects per study), we used robust variance estimation to adjust the standard errors and degrees of freedom for regression coefficients using the *clubSandwich* R package (Pustejovsky, 2018).

Primary Models

To examine the effects of intervention characteristics (or individual standards) on teacher and student outcomes, we estimated the following model, separately for each standard and outcome:

$$g_{jk} = \beta_0 + \beta_1 Standard_{jk} + \boldsymbol{X_{jk}} + u_k + e_{jk}$$

where g_{jk} is the effect size estimate j from study k, $Standard_{jk}$ is the standard (coded as 1 when the standard is included in the intervention, as described in the study; 0 otherwise); $\boldsymbol{X_{jk}}$ is the vector of methods control variable (publication status and level of random assignment for teacher outcomes; publication status, level of random assignment, and outcome measure type for student outcomes); u_k is a study-level random error term; and e_{jk} is the sampling error. β_1 represents the estimated effect of the standard on a given outcome.[4]

Moderator Models

To investigate sources of effect heterogeneity on teacher and student outcomes, we conducted moderator analyses for sets of intervention, sample, and outcome moderators. We ran a separate set of moderator analyses for each moderator grouping:

- Intervention feature: group training, instructional coaching, curriculum materials, in-person + online delivery, summer + school-year training
- Intervention content: instructional strategies, classroom management, focus on math, focus on reading, focus on science, instruction for ELs
- Intervention approach: instruction modified based on data, instructional skills practice during PD, teacher professional learning communities
- Sample characteristics: intervention length, intervention hours, years of experience, grade band
- Outcome: construct, content type, measure type

[4] The estimated meta-analytic mean effect is weighted by the inverse of the variance plus the estimated between-study variance, or $w_{jk} = 1/(v_{jk} + \hat{\tau}_u^2)$ (Borenstein et al., 2009).

We estimated the following model, separately for each moderator set and outcome:

$$g_{jk} = \beta_0 + \beta_m \boldsymbol{Moderators_{jk}} + \boldsymbol{X_{jk}} + u_k + e_{jk}$$

where $\boldsymbol{Moderators_{jk}}$ is the vector of moderators and all other variables are defined as above. β_m represents the estimated effect of moderator *m* on a given outcome.

Mediation Models

To investigate the extent to which changes in teacher instruction are associated with changes in student learning, we conducted a mediation analysis using the meta-analytic structural equation modeling (MASEM) approach outlined by Wilson et al. (2016). The approach includes a two-stage process: (1) collect and pool the individual studies' correlation matrices using meta-regression models with the same methods control variables as the primary model, and (2) conduct structural equation modeling (SEM). We used the *metaSEM* package in R (Cheung, 2014) to conduct the SEM analyses and estimate the corresponding paths in the theory of action. Specifically, we estimated the effects of the professional learning programs on teacher instruction (or outcomes) (path *a*), the effects of the professional learning programs on student outcomes (path *c'*), and the effects of teachers' instruction on student outcomes (path *b*) to extrapolate the mediating effect of teacher instruction on the effect of the professional learning programs on student outcomes (path *a*×*b*). We estimated the proportion of the change in teacher instruction to account for the total effect of professional learning on student achievement by dividing the mediation effect (path *a*×*b*) by the total effect (path *a*×*b* + path *c'*).

We conducted sensitivity analyses to ensure that the magnitude and direction of the effect for each path were consistent with model specifications. First, we limited the analysis to include only studies with effects of teacher instruction on student outcomes (path *b*).[5] Second, we removed two studies with particularly large effects from the analysis (i.e., Dolfin et al. [2019] and Gersten et al. [2010]). For both sensitivity analyses, the magnitude and direction of the path estimates were similar to those from the analysis with all studies.

[5] MASEM does not require complete data for all paths. Only 10 studies included data for path *b*. The main analysis included studies for which path *b* was missing. Thus, we conducted a sensitivity analysis on only the 10 studies for which data were available for all paths.

Appendix B. Included Studies and Evidence of the Standards

Table B.1. Summary Information of Studies Included in Meta-Analysis

Study	Number of Teachers	Number of Instruction Effects	Number of Students	Number of Achievement Effects	Number of 2011 Standards	Number of ADE	Number of 2022 Standards
Babinski et al. (2018)	30	4	105	7	3	3	7
Baker et al. (2015)	49	2	638	5	3	2	4
Blazar and Kraft (2015); Kraft and Blazar (2016)	79	6	N/A	N/A	2	1	2
Bos et al. (2012)	527	7	18,180	6	2	1	4
Brown et al. (2010)	82	4	N/A	N/A	5	4	8
Bruns et al. (2017)	3,121	6	50,463	4	3	2	2
Connor et al. (2011)	25	1	396	1	5	3	4
Cordray et al. (2012)	87	2	1,914	4	3	0	1
DeCesare et al. (2017)	77	10	1,189	6	1	1	2
Doabler (2010); Doabler et al. (2014)	129	10	N/A	N/A	0	1	3
Dolfin et al. (2019)	268	16	12,859	1	5	4	6
Faraclas (2018)	24	1	N/A	N/A	2	0	3
Garet et al. (2010); Garet et al. (2011)	179	3	4,528	9	2	2	4
Garet et al. (2017)	951	7	29,995	4	3	2	7
Garet et al. (2016)	165	6	1,697	2	4	2	6
Gersten et al. (2010)	81	2	468	6	2	1	5
Goldman et al. (2019)	48	6	964	2	4	1	5

Study	Number of Teachers	Number of Instruction Effects	Number of Students	Number of Achievement Effects	Number of 2011 Standards	Number of ADE	Number of 2022 Standards
Goodson et al. (2010)	128	4	1,296	3	2	2	3
Granger et al. (2012); Granger et al. (2019)	125	1	2,594	2	2	1	2
Grigg et al. (2013)	81	6	N/A	N/A	0	1	1
Jacob et al. (2017)	56	4	533	4	4	2	1
Jitendra et al. (2015)	82	11	1,898	3	1	1	2
Jitendra et al. (2016)	20	2	429	3	0	1	2
Johnson et al. (2019)	42	7	N/A	N/A	2	3	3
Johnson and Fargo (2010)	16	3	N/A	N/A	4	3	9
Jukes et al. (2017)	101	10	2,213	26	0	2	1
Kennedy et al. (2017)	12	4	N/A	N/A	2	2	2
Kim et al. (2011)	86	18	2,726	2	3	2	2
Lara-Alecio et al. (2018); Tang et al. (2020)	91	22	1,376	52	2	3	4
LaVenia (2016)	153	2	2,138	2	2	2	2
Matsumura et al. (2013)	167	1	2,983	2	4	3	6
Meyers et al. (2016)	84	6	1,964	4	7	4	6
Motoca et al. (2014)	138	22	N/A	N/A	4	2	1
Murray et al. (2018)	95	4	1,158	3	0	1	0
Nelson-Walker et al. (2013)	42	12	N/A	N/A	3	1	3
Nugent et al. (2016)	110	3	1,637	1	3	1	4
Ottmar et al. (2013); Rimm-Kaufman et al. (2014); Ottmar et al. (2015)	276	1	2,904	2	5	2	1
Parkinson et al. (2015)	130	2	4,333	5	3	3	3

Study	Number of Teachers	Number of Instruction Effects	Number of Students	Number of Achievement Effects	Number of 2011 Standards	Number of ADE	Number of 2022 Standards
Penuel et al. (2011)	27	3	836	3	2	1	2
Reddy et al. (2017); Fabiano et al. (2018)	89	4	N/A	N/A	3	1	4
Reinke et al. (2018)	104	1	1,680	2	2	2	3
Sailors and Price (2010)	44	2	444	1	2	1	3
Santagata et al. (2011)	44	2	N/A	N/A	3	2	3
Simmons (2010)	60	5	N/A	N/A	0	0	0
Supovitz (2013); Supovitz and Sirinides (2017)	64	4	1,347	3	6	1	3
Taylor et al. (2011); Taylor et al. (2013); Taylor et al. (2015)	53	1	3,052	1	3	1	3
Tong et al. (2019)	8	6	N/A	N/A	1	2	4
Vadasy et al. (2015)	61	30	1,232	5	1	1	4

Note. Number of teachers and number of students reflects the maximum number of teachers or students in a given study.

Table B.2. Evidence of Presence of Standards for Professional Learning and Areas for Deeper Exploration Across the 48 Included Studies

Standard	Number of Studies
2022 Standards for Professional Learning	
Conditions for Success	
Equity Foundations	8
Culture of Collaborative Inquiry	7
Leadership	1
Resources	12
Transformational Processes	
Equity Drivers	1
Evidence	18
Learning Designs	28
Implementation	35
Rigorous and Inclusive Content	
Equity Practices	4
Professional Expertise	32
Curriculum, Assessment, and Instruction	14
2011 Standards for Professional Learning	
Outcomes	25
Learning Designs	29
Data	9
Implementation	33
Learning Communities	9
Resources	11
Leadership	9
Areas for Deeper Exploration	
Equity of Cultural Competence	3
Content Expertise	28
Personalized Learning for Teachers	28
Personalized Learning for Students	1
Social and Emotional Learning Focus	5
Technology and Learning Innovation	9
Learning Systems	10

Suggested Citation: Garrett, R., Zhang, Q., Citkowicz, M., & Burr, L. (2021). *How Learning Forward's Professional Learning Standards are associated with teacher instruction and student achievement: A meta-analysis*. Washington, DC: Center on Great Teachers and Leaders at the American Institutes for Research.

Center on
GREAT TEACHERS
& LEADERS
at the American Institutes for Research®

GTLCENTER.ORG

AIR®
AMERICAN INSTITUTES FOR RESEARCH®

1400 Crystal Drive, 10th Floor
Arlington, VA 22202-3289
+1.202.403.5000 | AIR.ORG

15240_06/21