



RESEARCH REVIEW

Joellen Killion

Study offers keen insights into professional development research

► AT A GLANCE

Rethinking how to analyze and conduct research on professional development yields new insights to inform practice.

► THE STUDY

Kennedy, M. (2016). How does professional development improve teaching? *Review of Educational Research*, 86(4), 945-980.

•
Joellen Killion (joellen.killion@learningforward.org) is senior advisor to Learning Forward. In each issue of *The Learning Professional*, Killion explores a recent research study to help practitioners understand the impact of particular professional learning practices on student outcomes.

► WHAT THE STUDY SAYS

A new approach to analyzing professional development research provides both researchers and education practitioners useful information to guide their practice.

STUDY DESCRIPTION

Mary Kennedy conducts a review and analysis of the research on professional development in K-12 U.S. schools in the core content areas published since 1975.

Acknowledging that past reviews of professional development research based on its core features have insufficiently considered the variance in research designs and professional development content and design, Kennedy approaches the review with different theories of action about how professional development influences teacher learning and enactment of learning in practice.

The analysis yields a graphical as well as a statistical representation of effects that allows for alternative comparison of studies across contexts and for various types of interpretation. Kennedy's review sorts programs based on their theories of action. The theories of action include the core problems of practice the professional development addresses and the pedagogical approaches to teacher learning that supports and leads to enactment of learning.

QUESTIONS

Kennedy seeks to answer several questions.

- How do different professional development programs influence teacher learning?
- What problems of practice do professional development programs aim to address?
- What pedagogical approaches do professional development programs use to facilitate enactment or application of the content?
- What insights does a new approach to computing and displaying effect size of professional development studies between 1975 and 2014 that mitigates variances in studies of differing sample sizes, research designs, statistical analyses, and units of analysis offer researchers and practitioners?

METHODOLOGY

Kennedy conducted a search for experimental studies of professional development in core academic content areas (literacy, math, sciences, and social studies) within K-12 U.S. schools between 1975 and 2014. She established the review criteria as studies:

1. Of professional development only and excluded those with concomitant supports, such as curriculum or technology;
2. Conducted in the U.S.

“We need to ensure that PD promotes real learning rather than merely adding more noise to their working environment.”

— *Mary Kennedy*



only to accommodate the unique context of education, namely the lack of a national curriculum;

3. With evidence of student achievement either on distal measures of student achievement such as standardized assessments or state tests, coded as M1 outcomes, or proximal program-specific assessments of student achievement, coded as M2 outcomes;
4. With controls for teacher motivation to learn, namely voluntary participation versus mandatory participation;
5. With a minimum duration of one year; and

6. That follow teachers over time, rather than students.

The search yielded 28 studies to include in the analysis. Kennedy then designed a method for computing an estimate of program effects that accounted for sample size, unit of analysis, research design, and the study’s statistical procedures to minimize variance in effect sizes across studies.

ANALYSIS

Kennedy sorted the professional development programs included in the 28 studies based on two aspects of their theories of action, the program’s content and its approach to enacting the learning. The four content strands relate to the common problems of practice that challenge teachers:

portraying the content to students so that they can learn it; managing student behavior; enlisting student participation; and exposing students’ thinking to assess learning.

The second criterion for sorting professional development programs was the program’s approach to facilitating enactment — that is, the strategy the program employed to assist teachers in applying the ideas within their practice. Kennedy identified four approaches to enactment:

1. **Prescription**, which “explicitly describe or demonstrate ... the best way for teachers to address particular teaching problem” (p. 955) and with the expectation that teachers would follow the specific way with limited

► WHAT THIS MEANS FOR PRACTITIONERS

Kennedy addresses some basic assumptions about professional learning research, including its design and its focus on the common features of professional learning such as collective participation, content focus, intensity of duration, and learning designs such as professional learning communities and coaching.

Further, she questions previous computational approaches that fail to consider variance in the studies and that minimize the ability to compare studies. What her work offers practitioners are insights that relate to four of Learning Forward's Standards for Professional Learning (Learning Forward, 2011): professional learning's content (Outcomes standard); design (Learning Designs); approach to enactment (Implementation); and evaluation (Data).

Outcomes. Kennedy notes four areas related to common issues teachers experience in their classrooms and suggests that professional learning in any one area or all is likely to lead to increases in student achievement.

The Outcomes standard notes that the content of professional learning is related to student learning needs as defined by the content standards, educator learning needs as defined by their performance expectations, and programmatic or system needs as defined by strategic initiatives. The four areas of portraying content, managing student behavior, gaining participation, and exposing student thinking are common elements in teacher performance appraisal criteria.

Narrowing the focus of teacher professional learning to these four high-impact areas may be advisable and for leaders who are responsible for supporting teachers to gain expertise in these areas, especially, as Kennedy notes, in a

time when teachers face multiple competing demands. She states, "We need to ensure that PD promotes real learning rather than merely adding more noise to their working environment" (p. 974).

Learning Designs. This study, because it allows comparison across programs based on their approach to enactment, provides useful information to practitioners about the selection of learning designs and guidance for specific designs such as professional learning communities and coaching.

Kennedy urges researchers to "move past the concept of learning communities per se and begin examining the content such groups discuss and the nature of the intellectual work they are engaged in" (p. 972). The studies that included PLCs indicate that reading and engaging in facilitated discussions about the implications of research, for example, has a higher effect size than looking at students' achievement or their classroom practice without any guidance.

Coaching that uses a prescriptive approach has a lower effect size than coaching that uses the strategies or insights approach to enact learning. High-leverage designs for professional learning are the strategies and insights approach with prescription and body of knowledge having lesser effects.

Using common theories of action about how teachers learn and teacher motivation in professional learning program design and research can not only improve the results of professional learning, but also provide more useful information.

Implementation. The study suggests that multiyear programs have a greater effect size than those with a single year. While teachers don't have contact with the program in the second or third year necessarily, there is continued

2. **Strategies**, which is defining goals teachers seek to achieve and providing "a collection of illustrative practices that will achieve the goals" (p. 955);
3. **Insights**, which is "raising provocative questions that force teachers to re-examine familiar events and come to see them differently" so that teachers are making sound decisions and using professional judgment in classroom situations (p. 955);

4. **Body of knowledge**, which is developing "knowledge that is organized into a coherent body of interrelated concepts and principles and that can be summarized in books, diagrams, and lectures" and that gives teachers "maximum discretion regarding whether and how teachers would do anything with that knowledge" (p. 956).
Kennedy also considered how teachers were assigned to control

and treatment groups for each study. When assignment to groups was not comparable, such as mandatory for both treatment and control groups or voluntary for both, she excluded the studies to ensure commonality in teacher motivation to learn.

Kennedy organizes the 28 studies that meet the established criteria by the four common challenges teachers face. Fifteen studies address portraying content; two address student behavior; five address enlisting participation; and six address exposing student

following of the teachers' enactment of the learning and student achievement.

Measuring enactment and student achievement over time provides evidence that teacher learning is incremental and occurs over time. Learning Forward's Implementation standard calls for sustained, differentiated, classroom-based support over time to ensure enactment of learning. It also calls for ongoing, constructive feedback. Constructive feedback aligns with the strategies and insights approach to enactment of learning. For more information, see *The Feedback Process: Transforming Feedback for Professional Learning* (Learning Forward, 2015).

Data. The study calls for measuring enactment and student achievement over time. Learning Forward's Data standard calls for both formative and summative evaluation of professional learning using multiple forms and sources of data. This study suggests that the evaluation of professional learning occur over multiple years, a possible consideration for future revision of the standards for professional learning.

Other insights for practitioners include:

Motivation to learn. A prerequisite for professional learning, according to Learning Forward (2011), is "each educator involved in professional learning comes to the experience ready to learn" (p. 15). Comparing professional learning programs based on voluntary or mandatory assignment to treatment and control yields insights about the potential for negative effects not because of the quality of the learning experience, but rather because of learners' motivation to learn.

This study calls on practitioners to examine and address learners' motivation to learn for positive results and to reduce negative emotional effects that cause resistance or

resentment to the professional learning program. Kennedy reminds readers that attendance may be mandatory, yet learning is not. Future revisions of the standards might need to address learner motivation more explicitly.

Provider expertise and experience. Kennedy notes that studies with higher effect sizes are those whose providers have extensive practical experience and have expertise and experience in teacher learning, content, and pedagogy related to enacting learning. Providers' readiness, qualifications, and depth of expertise and experience influence the results of professional learning. Provider qualifications is another consideration for future revision of the standards.

Kennedy challenges some rudimentary assumptions long held in the field of professional learning and calls for actions that will both improve practice and the usefulness of research. "We need to replace our current conception of 'good' PD as comprising a collection of particular design features with a conception that is based on more nuanced understanding of what teachers do, what motivates them, and how they learn and grow. We also need to reconceptualize teachers as people with their own motivations and interests" (p. 974). As such, teachers deserve professional learning approaches that are intellectually rigorous about content meaningful to them rather than prescriptions and bodies of knowledge.

REFERENCE

Learning Forward. (2011). *Standards for Professional Learning*. Oxford, OH: Author.

thinking. She extends the summary by adding the approach to enacting the learning. Eight of the 28 studies use the prescription approach; 10 use strategies; seven use insights; and three use body of knowledge.

Kennedy displays the effects of each study in two graphical displays. In the first, using size, shapes, and color to denote descriptors of the studies such as its effects over time, sample size, type of outcome measure, unit of analysis, and post-professional development follow-up, she displays the effects of the

15 studies focused on the challenge of portraying curriculum content.

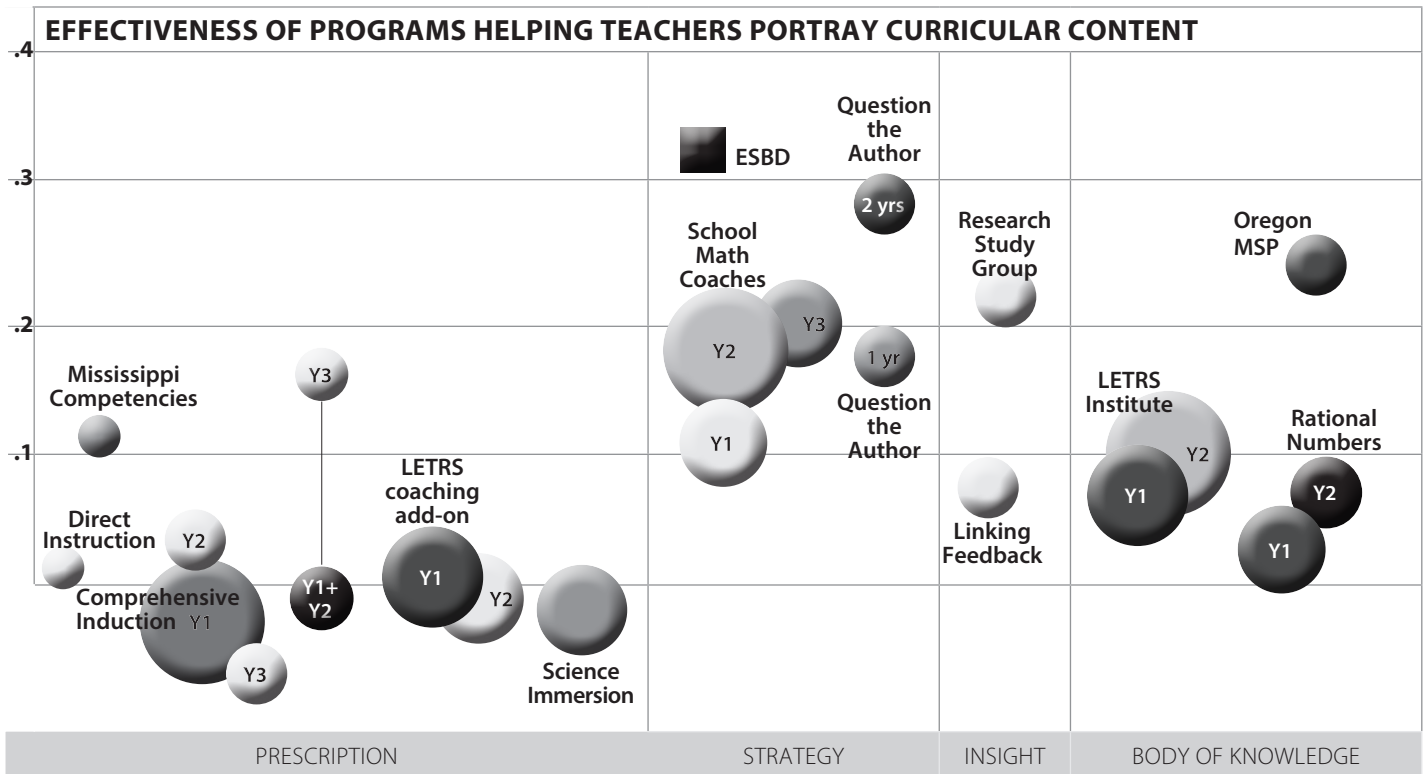
They are clustered together along the x-axis by their approach to enactment of learning, moving from prescription that limits teacher decision-making and judgment about enactment of learning to body of knowledge that provides the greatest amount of teacher choice to enact learning. Along the y-axis is the computed effect size. The graphical display, as Kennedy notes, makes it possible to compare programs within

and across content and approaches to enact learning.

In the second display, Kennedy uses the same size, shape, and color to denote the programs' descriptions to display the programs addressing each of the four challenge areas clustered together along the x-axis and their effect size on the y-axis.

RESULTS

Kennedy's computation and graphic display provide information about the programs that allows for comparison,



Source: Kennedy, M. (2016), *Review of Educational Research*, 86(4), 945-980. doi: 10.3102/003465431562. Reproduced with permission by Corwin.

interpretation, hypothesizing about interactions, and identifying implications for practitioners and researchers to consider.

For example, in the first display of programs focused on portraying curriculum content, programs with greater duration or program level of effort and state-of-the-art research design tend to show less effect on student achievement. Kennedy suggests this may be because of the prescriptive nature of the programs and their mandatory assignment. She hypothesizes that mandatory assignment, a trademark of most large-scale, high-duration programs, may reduce teachers' motivation to learn.

The overall display in the first figure depicts an inverted U-shape, suggesting that programs using strategy and insights as the enactment approach, those in the middle between prescription on the left and body of knowledge on the right, have a greater effect size than either of those

GUIDE TO READING THE GRAPHICAL DISPLAY

- Each icon in the graphic display portrays a unique program, clustered into the three different types of programs.
- Circles indicate programs with M1 measures of student achievement, such as standardized assessments that are more distal to the program goals; the square represents a program with an M2 measure of student achievement that is more specific or proximal to the program goals.
- The size of the icon indicates the number of teachers involved in the program. For example, School Math Coaches had 418 participating teachers compared to Linking Feedback that had 81.
- When the program occurred over multiple years, each year of the program is displayed and indicates as Year 1, Year 2, and Year 3, as in the case of School Math Coaches.
- The darkness is a representation of the program intensity as determined by the amount of time teachers engaged in professional development. Linking Feedback involved three hours of professional development compared to 67 hours for Rational Numbers.
- The icons closer to the top of the display have larger effect sizes than those at the bottom. The display provides opportunities to compare and hypothesize about the effects of different types of professional development programs.

categories.

Body of knowledge enactment has a higher effect size than prescription. Programs that had multiyear follow-up tend to have higher effect sizes than

programs without follow-up.

In this latter group of studies, teachers did not necessarily have contact with the program after year one, yet their enactment of learning

and student achievement was followed. If coaching is included, coaching that emphasizes strategy and insights tends to be more successful than prescription-oriented coaching.

Kennedy notes that teacher practice, as previous research confirms, increases incrementally over time. The display also confirms that M2 outcome measures, those more closely aligned to program content and goals, have greater effect sizes than those that use more general measures of student achievement.

Kennedy summarizes the first display by noting that prescription as an approach to enactment has the lowest effect, with body of knowledge next, insight next, and strategy the highest for the studies addressing the challenge of portraying content.

Studies with mandatory assignment have lower effect sizes as do larger-scale studies than other studies. The overall effect size is .10 for these 15 studies and, when the studies that used mandatory assignment are removed, the effect size rises to .16.

The studies in this cluster are all below .2, and Kennedy notes that higher effect sizes in other reviews are likely distorted by the variance in the sample size, research design, statistical analysis, and professional development content and approach.

The second display includes all the programs clustered by the challenges their content addresses. Using the same symbols to depict each program, clustered along the x-axis by their challenge area and excluding those programs with mandatory assignment in the portraying content section, Kennedy makes it easy to compare programs based on participant assignment to treatment group.

The differences in effect size introduce the possibility, Kennedy suggests, that social motivation, in which the participants desire to support the researcher rather than perceive a

Mary Kennedy emphasizes that programs that acknowledge the incremental growth of teachers and include a follow-up measure have larger effect sizes.

need to improve their practice or learn something new, may be at play in instances where effect size is larger and where there is follow-up with teachers.

She specifically points to programs that had a negative effect, and posits that such a negative emotional response may be resistance to the program's demands. Programs in the areas of enlisting participation and exposing student thinking tended to be more strategy- and insight-based and have larger effect sizes, especially in their second year, than other programs.

In concluding about the second display, Kennedy notes that programs in any of the four challenge areas are likely to improve student achievement, suggesting that no one area is more important than another. All contribute to improved practice and student success.

Kennedy explains how the approach she used for analysis of the effect size of the 28 studies differs from the more traditional analysis of professional development studies using the common features of intensity, collective participation, content knowledge, and coaching.

She challenges basic assumptions in each area with studies she included and calls upon researchers to go beyond the surface features to examine more closely the specific content of and approach to enactment using the theories of action she articulates and teacher

motivation to learn as significant factors in the success of professional development. She also notes that how coaching, a relatively common feature in professional learning today, supports enactment influences the effect size.

She calls on researchers to examine more closely professional development providers' content and pedagogical knowledge and their approach to enacting learning. She notes that more effective programs included in the 28 studies had providers with established histories of working with teachers, direct experience in the classroom, and expertise with the content and teacher learning.

She emphasizes that programs that acknowledge the incremental growth of teachers and include a follow-up measure have larger effect sizes.

LIMITATIONS

Kennedy introduces a new way to analyze the effects of professional development research that challenges the What Works Clearinghouse standards for research design and demonstrates that research in professional development that follows the recommended high-level evidence standards are less effective than studies that are smaller scale and use voluntary assignment. As she notes, other factors not examined in the traditional professional development research, such as motivation to learn and provider attributes, may influence results.

Kennedy does not include the specific effect size for each study based on her computation. The effect size is portrayed in the graphical display, yet including the specific number in the table would be a helpful reference for readers. Overall, the computation of effect size produces small numbers for each study, which may lead some to question the value of professional development in general based on the small effect sizes. ■