Is your professional learning working?
8 steps to find out

BY JOELLEN KILLION

Evaluating professional learning to measure its effectiveness and impact on student learning is an important process for those who design, lead, and facilitate educator learning. Evaluation demonstrates a commitment to accountability for investments in professional learning and a mechanism to ensure its continuous improvement. The Standards for Professional Learning (Learning Forward, 2022) articulate the importance of evaluation. According to the Evidence standard, "Professional learning results in equitable and excellent outcomes for all students when educators create expectations and build capacity for use of evidence, leverage evidence, data, and research from multiple sources to plan educator learning, and measure and report the impact of professional learning" (Learning Forward, 2022).

Evaluating professional learning requires thoughtful and intentional effort. As I have defined it in *Assessing Impact* (Killion, 2018), "Evaluation is a systemic, purposeful, standards-driven process of studying, reviewing, and analyzing data about a professional learning program gathered from multiple sources to make judgments and informed decisions about the program" (Killion, 2018, p. 8).

ABOUT THIS ARTICLE

In fall 2003, Learning Forward (then National Staff Development Council) published the article "8 smooth steps" by Joellen Killion, which outlined an eight-step process for evaluating a professional learning program. Since then, the article has become a model for designing effective evaluation of professional learning. In this article, Killion offers a fresh look at how to apply a scientific, systematic process to evaluation that ensures reliable, valid results.

TOOLS ACCOMPANYING THIS ARTICLE

- Mapping an evaluation step by step
- Identifying KASABs
- Creating a logic model
- Establishing an evaluation framework

OVERVIEW OF STEPS

1. Assess evaluability.
2. Formulate evaluation questions.
3. Construct evaluation framework.
4. Collect data.
5. Organize, analyze, and display data.
6. Interpret data.
7. Report, disseminate, and use findings.
8. Evaluate the evaluation.

However, not all evaluation efforts are rigorous enough or sufficient to make a claim that professional learning impacts student learning.

Using the following eight steps of the evaluation process, drawn from extensive practice and research in program evaluation, professional learning leaders will be able to answer important questions about the relationship among professional learning, educator practice, and student learning. The process is described in this article. Threaded throughout is an example of evaluating a professional learning effort to improve teachers' math instruction to improve student achievement in math.

1. ASSESS EVALUABILITY.

The first step is determining whether a professional learning program or initiative is ready to be evaluated. This is based on the degree to which the professional learning, as planned, is sufficient to generate the intended results. A program of professional learning, not individual events or episodes, is far more likely to change educator practice and student learning. A program of professional learning is a "set of planned and implemented actions, guided by research, evidence, and standards of effective professional learning, accompanied by adequate resources, and directed toward the achievement of defined outcomes related to educator practice and its impact on student learning" (Killion, 2018, p. 8).

Assessing evaluability involves determining if the program's design is likely to produce its intended results. "It is futile to expect results for students from a professional learning program that is unlikely to produce them. Evaluation cannot compensate for a professional learning program that is poorly conceived and constructed. Perhaps Chen (1990) said it best: 'Current problems and limitations of program evaluation lie more with lack of adequate conceptual framework of the program than with methodological weakness (p. 293)' " (Killion, 2018, pp. 44-45).

Before evaluating any professional learning program, the evaluator asks whether the program is feasible, clear, sufficiently powerful to produce the intended results, and worth doing. To determine whether a program is ready to be evaluated, an evaluator analyzes the program's goals, expressed as expected changes for students; outcomes, expressed as changes for educators; indicators of success; standard of success; and the program's theory of change and logic model, each of which is described below.

Goals

A program's goals express its intended results in terms of student success. Instead of a goal such as training all teachers, a results-driven program has a clearly stated goal for students, such as all students meeting grade-level expectations in math. A strong goal is to increase student performance on end-of-course assessments by a certain percentage over a defined time period. When student performance reaches the established threshold, the program is working as intended. If not, the program requires adjustment in its design, operations, or resources.

Outcomes

Outcomes describe the specific changes necessary to achieve the goal. The changes occur in some or all of these areas: knowledge, attitudes, skills, aspirations, and behaviors, which can be remembered through the acronym

KASAB. These changes occur in practices associated with the student goal area and, for professional learning, they are most closely connected to changes educators make. They can occur among multiple stakeholders. For example, in some cases, coaches are expected to change their practices to support desired changes in teachers' instructional practices. Principals, too, will be required to change some of their monitoring and supervisory practices to align with the classroom changes. When principals change their supervisory practices, central office may make changes to help principals spend more time in classrooms. Even parents may be expected to make changes, if they are asked to modify how they are supporting students at home.

Indicator of success

An indicator of success is the specific way success is demonstrated. Any goal and set of outcomes might have multiple indicators of success to strengthen the reliability of the claims evaluators are able to make about the program. Professional learning designers identify one or more indicators of success early in the planning process so the program's design aligns with the expected results. An indicator of success also guides an evaluator to know what type of data to collect. Indicators can include formative assessments, classroom tasks, observations, enrollment of underrepresented populations in advanced-level courses, grades, and performance on national standardized tests.

Standard of success

A program's standard of success is the benchmark that defines whether the program has achieved its goals. It typically is a number representing the performance increase that, when met, is sufficient to declare the program successful (or, when not met, to signal the need for adjustments). For example, a 10% increase in math performance on a common formative assessment is a standard of success. Specificity is important; if evaluators have not set a specific target, then any degree of improvement, even a 0.002 increase in average test scores, may be interpreted as success, even if it is not practically meaningful. For example, for changes in educator practice associated with a mathematics professional learning program, a standard of success is setting the expected level for teachers' accuracy and frequency of implementation of the mathematical instructional practices at 85% during year one and increasing it gradually to 100% by the third year.

Theory of change

A theory of change articulates "what the professional learning program is and how it is expected to produce the intended results. A program's theory of change delineates the causal processes through which change happens as a result of the program's strategies or actions" (Killion, 2018, p. 54). It includes the program's components, their sequence, and the assumptions upon which the program is based (Killion, 2018). An explicit theory of change is a road map for program designers, managers, participants, and evaluators showing how the program will work and how they see the connection between educator learning and student success. It is the big picture that serves as a planning tool, an implementation guide, a monitoring tool, and a tool for evaluating the program's success. Without the theory of change, the connection between the program's components and intended results — especially the connection between educator learning and student improvement — may be unclear.

Theories of change can be based on existing theories, research, or best practice. For example, the social interaction theory of learning might serve as the basis for designing adult learning, in which case the theory of change would include multiple, frequent, in-depth opportunities for participants to process their learning with colleagues.

Consider the sample theory of change on p. 61 for a mathematics professional learning program.

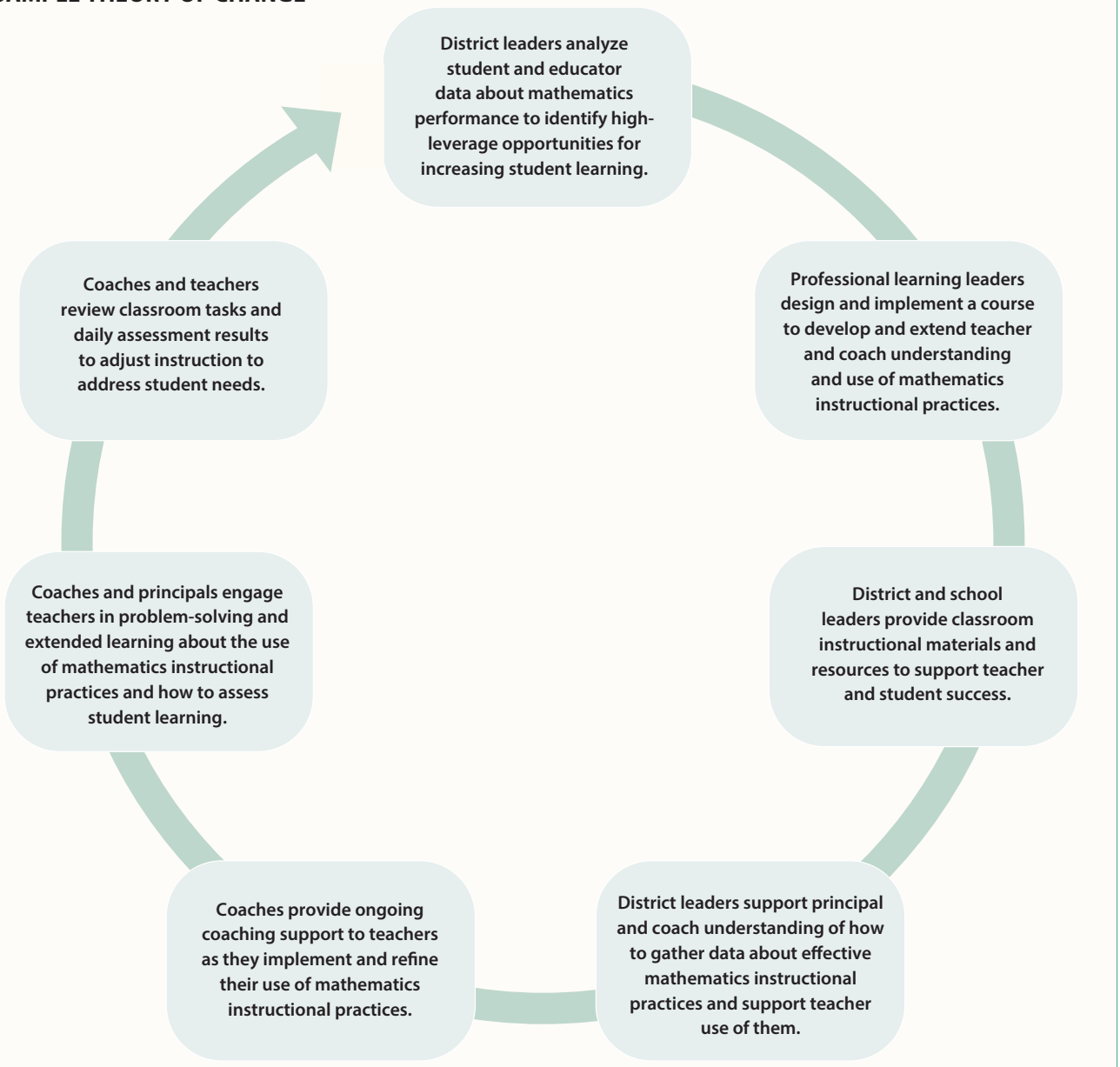
Every theory of change is based on a set of assumptions that guide decisions about the components included and their sequence. For example, here is one assumption on which this theory of change is based: Coaching enhances the implementation of instructional practices.

Any one program can have multiple theories of change. Individual theories are neither right nor wrong, but one may be more appropriate for a specific context and circumstances. Furthermore, when multiple actors are expected to influence the success of professional learning, there may be several theories of change, each related to a specific group of stakeholders, such as central office staff, principals, and coaches. Those theories of change layer together, as depicted on p. 62.

Logic model

A logic model is a particular kind of action plan that specifies the inputs, activities, outputs, and initial, intermediate, and intended outcomes that will accomplish the identified goal. A program's logic model is related to but distinct from the theory of change: "A theory of change identifies rationale for the chain of causal actions that predicts and explains how the program works to achieve the intended results. … A logic model uses the theory of change to depict the operation of a program by delineating several key components of an action or operational plan" (Killion, 2018, p. 60). A logic model ensures that all the program's activities align with the intended outcomes and that initial and intermediate outcomes will lead to the intended results. An evaluator uses the logic model to assess the thoroughness of the plan before beginning an evaluation and uses it as a progress map in the program's formative evaluation, which focuses on implementation of the program and benchmarks that lead to goal attainment.

SAMPLE THEORY OF CHANGE



A logic model has several components:

- Inputs: Resources assigned to a program, including personnel, facilities, equipment, budget, etc.
- Outputs: Products generated to support program implementation or documentation of a completed action.
- Activities: Services the program provides to clients.
- Initial outcomes: First-level changes in program participant knowledge and skills as a result of early activities. They may include changes in attitude. They have little inherent value, yet are important precursors to later-stage changes.
- Intermediate outcomes: Changes in program participant attitudes, aspirations, and behaviors resulting from the initial changes in knowledge and skills. They are essential to lead to changes in classroom practices that affect student learning experiences and success.
- Intended results: Desired results of the program (related to student achievement) expressed in the goal.

For an example, see the logic model on p. 63.

KASAB

Knowledge
Conceptual understanding of information, theories, principles, and research.

Attitudes
Beliefs about the value of information or strategies.

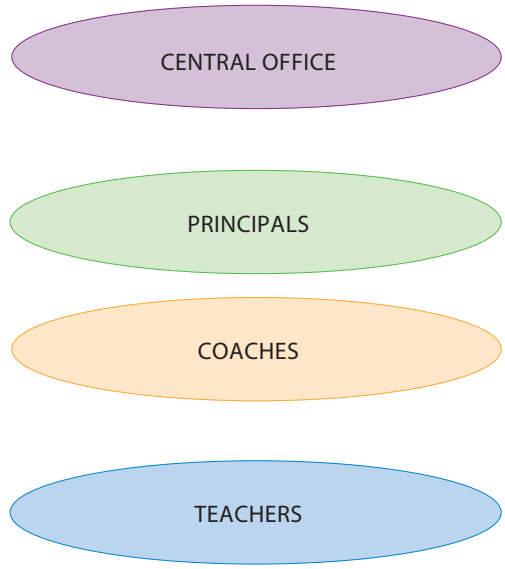
Skills
The ability or capacity to use strategies and processes to apply knowledge.

Aspirations
Desires, or internal motivation, to engage in a practice.

Behaviors
Consistent application of knowledge and skills driven by attitudes and aspirations.

Source: Killion, 2018, p. 50.

HOW THEORIES OF CHANGE LAYER TOGETHER



Logic models are planned from the goals backward. The intended results for students drive the intermediate and initial outcomes, which in turn drive the actions teachers are expected to take and changes they are expected to make. These actions drive the outputs that will be created to support the changes and the inputs necessary to produce them. This backward planning is represented in the sample logic model by arrows at the top of the figure.

2. FORMULATE EVALUATION QUESTIONS.

Evaluators craft formative and summative evaluation questions that allow them to make claims about the effects of the professional learning program. The professional learning goals, outcomes, theory of change, and the logic model are used to generate the evaluation questions.

Formative evaluation questions are based on the program's initial and intermediate outcomes. Without first answering these questions, evaluators will be unable to claim that teachers' learning contributes to student learning. Most professional learning programs have numerous initial and intermediate outcomes, so evaluators choose which are pivotal to the program's success and most likely to provide crucial information about whether the program is likely to lead to student improvement. For example, for the theory of change and logic model in this article, an evaluator may choose to focus on whether teachers are integrating the new practices in their instruction, rather than whether they can explain the practices, because use of the practices is more critical in student success. They may also decide to examine whether coaches and principals are conducting classroom observations of math instruction and using the data collected in reflective conversations with teachers.

Answering these kinds of formative evaluation questions helps program leaders monitor progress toward the goal so that they can adjust the program design, if necessary, to ensure that the actions are leading toward the goal. It also helps them explain how the change steps are connected, which lends more support to claims about how the program impacts student achievement. For example, if the evaluation shows that teachers engaged in coaching to support the implementation of specific practices, teachers then implemented the practices with accuracy and frequency, and student achievement increased, the formative data will help make the case that the professional learning positively influences achievement.

Summative evaluation questions ask whether the program met its goals. A summative evaluation question for the goal above is: Does student success on end-of-unit and end-of-course assessments increase by 10% for all students when teachers are integrating the mathematics practices accurately and routinely in instruction? It is important for the goal and summative questions to examine student achievement. If the goals are not expressed as student success goals, then the evaluator may be able

SAMPLE LOGIC MODEL

Goal statement (intended results for students: **By the end of the school year, all students will increase their performance on end-of-unit and end-of-course assessments in mathematics by at least 10%.**

Order of planning: Logic models are planned from the goals backward. ← ← ←

INPUTS		OUTPUTS	OUTCOMES		RESULTS
Resources	Actions/strategies	Outputs	Initial educator learning outcomes	Intermediate educator practice outcomes	Intended results for students
Time to participate in the professional learning course. Math manipulatives.	Design professional learning course for teachers, coaches, and principals. Implement course.	Course materials. Video library of teachers applying the mathematics practices.	Teachers, coaches, and principals explain the eight mathematics practices with 85% accuracy (*knowledge*). Teachers, coaches, and principals identify the research-based explanation about how students benefit when mathematics practices are used (*knowledge, attitudes*). Teachers, coaches, and principals evaluate videotaped lessons for evidence of integration of mathematics practices with 85% accuracy (*knowledge, skills*).	All teachers and coaches remodel instructional units to integrate mathematics practices and explain the rationale for choosing the practices included and their alignment with the unit's content with 85% accuracy (*skills*).	All students increase their performance on end-of-course assessment by at least 10% by the end of the school year.
Math instructional materials that integrate the mathematic principles.	Engage teachers in analysis of instructional materials for integration of mathematics practices.	Formative assessments of student performance.	Teachers explain how to assess student performance that results from use of the mathematics practices with 85% accuracy (*knowledge, skills*).	Teachers generate daily and common formative assessments aligned with mathematics content standards and the mathematics practices.	
Time for microteaching.	Engage teachers in microteaching to apply the mathematics practices.	Classroom observation guide for implementation of the mathematics practices. Innovation Configuration map for self-assessment of the implementation of mathematics practices.	Teachers, coaches, and principals evaluate videotaped lessons for evidence of integration of mathematics practices with 85% accuracy (*knowledge, skills*).	Teachers, coaches, and principals explain the value of and advocate for the use of mathematics practices to colleagues, students, parents, and community members (*attitudes*).	
Time for teacher teams to plan and design units and lessons and revise student tasks, discuss challenges, and extend their content and pedagogical understanding.	Engage teachers and coaches in planning for instruction using the mathematics practices. Model instruction that integrates mathematics practices. Implement units and lessons integrating the mathematics practices.	Unit and lesson plan accommodations to meet needs of diverse learners. Criteria checklist for rating the integration of mathematics practices in planned and taught lessons and units.	Teachers generate a unit of instruction that integrates at least four of the mathematics practices that meet 90% of the criteria in the integration of practices checklist. All principals and coaches acquire data-gathering strategies to use in classroom observation about the use of the mathematics practices (*skills*).	Teachers generate five units of instruction that integrate at least four math practices with 80% accuracy. Teachers implement at least four of the eight mathematics practices into a unit of instruction with 80% accuracy five times during the school year (*behaviors*). Teachers use the criteria checklist for integration of mathematics practices to analyze their instructional lessons, student work products, and formative assessment data to reflect on the strengths of their application of the mathematics practices and identify opportunities for refining future instruction (*behaviors, attitudes, aspirations*). Teachers increase the frequency of their use of appropriate mathematics practices to 95% in all math instruction (*behaviors*). All principals and coaches acquire data-gathering strategies to use in classroom observation about the use of the mathematics practices (*skills*).	

to make claims about the degree to which the program achieved the initial or intermediate outcomes, but not its impact on student learning.

3. CONSTRUCT EVALUATION FRAMEWORK.

The evaluation framework is the plan for the evaluation. Decisions made in this step determine the type of data necessary to answer the formative and summative evaluation questions, the appropriate sources of those data, appropriate and feasible data collection methods, data analysis processes, timeline for data collection and analysis, and responsible persons.

These decisions influence the reliability and validity of claims made using the data, so evaluators seek the most robust data possible, along with ways to triangulate data sources, types of data, or data collection methods to strengthen the claims. In some cases, though, evaluators may decide to use approximate data because of feasibility issues. If, for example, evaluators want to know whether teachers are implementing mathematics practices, direct observations of classrooms is the most authentic source of data to answer this question. However, observing every teacher for the purpose of program evaluation is a costly and labor-intensive process. The evaluator might therefore use a teacher survey about how often they use the practices. Because this is an approximate data source, the evaluator may want to supplement the survey data with coaching notes and student work samples to triangulate the data.

In addition to the types of data, the source of the data matters. Ideally, a matched group of students and/or teachers not receiving the program is available for comparisons, yet that is not always feasible in practitioner-based evaluations. In this case, an evaluator may use a pre- and post-program design because two points of data are required to answer a question about an increase either in educator practice or student success.

For example, see the sample elements for an evaluation plan on p. 65. Note that this is not a comprehensive evaluation plan, but rather an excerpt for illustrative purposes.

4. COLLECT DATA.

Data collection requires a systematic and thoughtful process to ensure that data are accurate. To ensure accuracy in this step, evaluators often create checks and balances to ensure that data are recorded precisely, errors in data entry are found and corrected, and missing data or outlier data are handled appropriately. Sloppy data collection and management can compromise the integrity of even the most well-designed evaluation.

When collecting data from human subjects, evaluators adhere to standards established by the Joint Committee on Standards for Educational Evaluation (Yarbrough et al., 2010). They ensure that they have met all the policy expectations of schools and districts for notification, privacy of records, or other areas, and abide by the evaluator code of ethics and standards.

5. ORGANIZE, ANALYZE, AND DISPLAY DATA.

"Throughout the data analysis process, the evaluator is constantly looking at new ways to combine, unpack, rearrange, and connect data to understand the program being evaluated" (Killion, 2018, p. 135). As data are collected, evaluators organize it, check its accuracy, and prepare for analysis. Evaluators pilot newly developed or modified data collection instruments to check the instruments' accuracy and clarity. If more than one individual is collecting data, data collectors may calibrate their processes to achieve accuracy and consistency. Evaluators check for any abnormalities in the data set such as inaccuracies or incompleteness in recording data.

Once evaluators are confident in the integrity of the data, they analyze it. Many practitioners hesitate to use inferential statistical analyses, yet in most cases descriptive analyses such as counting totals, finding patterns and trends, or simple calculations such as determining the mean, median, mode, and range are sufficient to answer most evaluation questions. Some evaluation questions may require more sophisticated analyses such as factoring, assessing covariance, or creating statistical modeling.

After analyzing data, evaluators display the analyzed data in charts, tables, graphs, or other appropriate formats for interpretation. Careful titling and labeling of data displays facilitate data interpretation and make it more useful for stakeholders.

6. INTERPRET DATA.

While data analysis is the process of counting and comparing, interpreting is making sense of what the analysis tells us. It is a collaborative process carried out by program designers and diverse key stakeholders, including participants, who bring different perspectives into the process. In most evaluations of professional learning programs, this means that teachers, principals, central office staff, and sometimes students work together to study the analyzed and displayed data and form claims about the program's effectiveness and impact on student learning.

Interpretation involves three parts: making meaning, which is the process of determining the significance and explanation of the findings; judgment, which brings values to bear to determine merit and worth; and recommendations, which propose actions based on the results (Patton, 2008). For example, if the analysis demonstrates that math scores have gone up over three years, in the interpretation phase, evaluators engage stakeholders in exploring what the increase means in terms of the professional learning program, considering questions such as: What contributed to the increase? Was the increase consistent across all grades and student populations? What does the increase mean for our school's

SAMPLE ELEMENTS OF AN EVALUATION FRAMEWORK

Professional learning program goal: **By the end of the school year, all students will increase their performance on end-of-unit and end-of-course assessments in mathematics by at least 10%.**

Measurable outcomes/ changes	Evaluation questions (formative and summative)	Data / evidence needed	Data source	Data collection method	Data analysis method	Timeline	Responsible person(s)
Goal: All students increase their performance on end-of-course assessment by at least 10% by the end of the school year.	Did all students increase their performance on end-of-course assessment by at least 10% by the end of the school year?	Student performance data on end-of-course assessment.	Students.	Assessment scores.	Comparison with previous year's student performance.	May–May.	Assistant principal; math teachers.
Teachers generate five units of instruction that integrate at least four math practices with 80% accuracy.	Did teachers' five units of instruction integrate at least four mathematics practices with 80% accuracy?	Teacher work products — generated units.	Teachers.	Collect units monthly November through April.	Scoring unit plans with criteria checklist.	November, January, February, March, and April.	Assistant principal, coach, district math specialist.
Teachers implement at least four of the eight mathematics practices into a unit of instruction with 80% accuracy five times during the school year.	Did teachers implement at least four mathematics practices into a unit of instruction with 80% accuracy at least five times during the school year?	Teacher classroom practice data.	Teachers. Student work products.	Classroom observation data. Teacher self-assessment using the Innovation Configuration map. Teacher reflection notes.	Trend analysis of practices (number and type in observed class) evident in observation and reflection notes. Means score on teacher self-assessment.	Monthly, November through April.	Assistant principal, coach, district math specialist.

math instruction going forward? Then stakeholders make a judgment about the program's merit and worth and recommend its continuation, modification, or discontinuation.

During the interpretation phase, claims of contribution — those stating that the program may have or likely influenced student success — can be made when the evaluation design is descriptive or quasi-experimental. But claims of attribution — that professional learning was the definite cause of the results — require experimental, randomized designs, which are not often used in practitioner-led evaluation studies.

7. REPORT, DISSEMINATE, AND USE FINDINGS.

After interpretation, evaluators decide the audiences to whom they will report the results and the most appropriate formats in which to share them. Not all audiences want the same kind of report. Some formats for sharing evaluation results include technical reports, brief executive summaries, pamphlets, newsletters, news releases to local media, and presentations.

A significant benefit of professional learning program evaluations is using lessons learned to improve future programs. This is most likely to happen when program evaluation results are widely shared, discussed, and used. Evaluators share in the responsibility with program managers or leaders

to use the results of an evaluation in making decisions about subsequent programs or modifications to the existing one.

Evaluation of professional learning increases the integrity of the field and contributes to its body of knowledge. "As the field of professional learning continues to grow and establish a firmer foundation based on evidence, every evaluation offers an opportunity to contribute knowledge to that foundation. Building on the successes of current professional learning and avoiding, when possible, past challenges, professional learning leaders will be able to design, implement, and evaluate more professional learning that has greater positive effects on educators and their students" (Killion, 2018, p. 194).

8. EVALUATE THE EVALUATION.

Evaluations rarely include this step, which is a missed opportunity. Evaluating the evaluation, a meta-evaluation, involves reflecting on the evaluation processes (rather than its results) to assess the evaluator's decisions and skillfulness, tools and processes used, the resources expended for evaluation, and the overall effectiveness of the evaluation process. Evaluating the process is an opportunity to improve future evaluations and strengthen evaluators' practice. It may include critical friends who can review and reflect with the evaluator about how the evaluation was done. "When evaluators seek to improve their own work, increase the use of evaluation within an organization, and build the capacity of others to engage in evaluation think, they contribute to a greater purpose. Through their work, they convey the importance of evaluation as a process for improvement and ultimately for increasing the focus on results" (Killion, 2018, p. 200). This process is the hallmark of a reflective practitioner and, like all professionals, evaluators commit to continuous improvement by examining the effects of their decisions and actions.

INVESTING IN STUDENT SUCCESS

Evaluating professional learning requires applying a scientific, systematic process to ensure reliable, valid results. The effort required can be significant, but it is worthwhile. Evaluation not only provides information to determine whether professional learning impacts student success, it also provides information about how to strengthen efforts to increase the potential for future success. Professional learning leaders face challenging decisions about how they invest their resources and effort to ensure that they have the greatest potential for increasing student success, and evaluations can provide the evidence needed to make these critical decisions, be accountable and responsible for investments, and contribute to strengthening the field of professional learning in the future.

REFERENCES

Chen, H. (1990). *Theory-driven evaluation*. Sage Publications.

Killion, J. (2018). *Assessing impact: Evaluating professional learning* (3rd ed.). Corwin & Learning Forward.

Learning Forward (2022). *Standards for Professional Learning*. Author.

Patton, M.Q. (2008). *Utilization-focused evaluation* (4th ed.). Sage.

Yarbrough, D.B., Shula, L.M., Hopson, R.K., & Caruthers, F.A. (2010). *The Program Evaluation Standards: A guide for evaluators and evaluation users* (3rd ed.). Corwin Press.

•

Joellen Killion (joellen.killion@learningforward.org) is a senior advisor to Learning Forward.

Mapping an evaluation step by step →

This tool provides guiding questions to plan and conduct a professional learning evaluation, using the eight steps of evaluation described on the previous pages. In collaboration with other stakeholders (e.g., professional learning designers, participants, and evaluators), discuss the questions and take detailed notes.

To begin, ask:

- What is the purpose of this evaluation?
- Who are the primary users/stakeholders of the evaluation results?
- What is their intended plan for using the results?

Then, use the following questions to map out each of the eight evaluation steps. The tables on pp. 69-70 provide spaces to articulate the desired changes in knowledge, attitudes, skills, actions, and beliefs and your logic model.

Planning an evaluation is a time-intensive process. These questions and steps will likely not be completed all at once.

1. Assess evaluability. Evaluators review the professional learning plan to make sure it is ready to be evaluated and, if needed, work with stakeholders to make changes to ensure the greatest likelihood for program success.	1. What are the professional learning program's goals? • Are they student-focused and results-oriented? • Are they measurable, time-bound, equitable, and inclusive? 2. What are the professional learning program's intended outcomes for educators? • Are they measurable, time-bound, and results-oriented? • Do they specify the intended change (knowledge, attitudes, skills, aspirations, behaviors)? • Are they plausible and focused on educator behaviors/practices? 3. Have the indicators of success and standards for success been set for all outcomes? 4. What is the professional learning program's theory of change and the assumptions upon which it is based? Has it been reviewed by representative program stakeholders and participants? 5. Is the professional learning program's logic model complete? In other words, what are the inputs, activities, initial outcomes, intermediate outcomes, and intended results of this program? 6. Based on the status of the professional learning program plan, is this evaluation ready to initiate, or are adjustments in the program design needed first?
2. Formulate evaluation questions. Evaluators use the goals of the professional learning program to write the formative and summative evaluation questions that will drive the evaluation.	1. What are the evaluation questions? • Formative • Summative 2. How well do the evaluation questions reflect the interests of the primary stakeholders? 3. How well do the evaluation questions align with the program's goals and the evaluation's purpose? 4. Are the evaluation questions: • Reasonable, appropriate, and answerable? • Specific about success indicators? • Specific about the measure of program success?
3. Construct evaluation framework. Evaluators plan how to answer the evaluation questions, deciding what data to collect, from whom, how, and when, and how to analyze the data once they are collected.	1. Who will conduct the evaluation — a stakeholder internal to the program or system, an external evaluator (e.g., from a research organization), or a combination? 2. How will the evaluation question(s) be answered? • What are the key constructs/variables that will be measured? How have key terms (such as student achievement, improvement, increase, and professional learning) been defined so that they are clear and specific and aligned with the indicators of success? • What type of evaluation design is needed to answer the evaluation questions? Do the questions require making a comparison to determine impact? If so, what are possible comparison groups? Which is the most appropriate comparison group for this evaluation? 3. What will the data plan be? • What kind of data can provide evidence that the intended changes have occurred? • Who are the data sources that will provide evidence of the intended change? How essential is it to have multiple data sources for this evaluation? • What data collection methodologies are most appropriate to obtain the needed data? • When and where will the data be collected? • How will data be analyzed? 4. How much will the evaluation cost? • Are resources, including time, fiscal resources, and personnel, available to conduct this evaluation? • If resources are not adequate, what aspects of the evaluation plan can be modified without compromising the integrity of the evaluation? • Is the evaluation worth doing given the cost and potential modifications? 5. Who is responsible for each part of the evaluation? 6. Have primary stakeholders reviewed and approved the evaluation plan?

continued on next page

4. Collect data. To begin the work of collecting data, evaluators prepare and field-test instruments, calibrate scoring, establish processes for managing data, and determine how to address missing or erroneous data.	1. Have the instruments and procedures for data collection been field-tested? 2. What revisions are necessary? 3. How will data collectors be trained? 4. After early data collection, do any data seem redundant? What are the advantages and disadvantages of continuing to collect these data? Is it appropriate to continue or discontinue collecting these data? 5. After early data collection, what data seem to be missing? Is it essential to collect these missing data? If so, how will a new data collection methodology be implemented to collect these data? 6. What processes have been established to manage data collection and recording? 7. What processes have been established to ensure safekeeping and integrity of data?
5. Organize, analyze, and display data. With data in hand, evaluators organize the data, analyze it using predetermined descriptive or inferential statistical procedures, display the analyzed data, and formulate findings from the analyzed data.	1. How will data be sorted, grouped, or arranged before analysis? 2. How will missing data be handled in statistical analyses? 3. How will data be displayed to facilitate interpretation and understanding? 4. How clearly and succinctly are the data findings stated?
6. Interpret data. This step engages stakeholders in interpreting the analyzed data and findings to make and support claims and recommendations based on the analyzed data.	1. What do these data mean? 2. What interpretations and claims can be made from these data? 3. How well-supported are the interpretations and claims? 4. Have possible alternative interpretations been considered? 5. Does this evaluation support claims of attribution or contribution? 6. Does this program have merit, worth, and significance? 7. What recommended actions can help program stakeholders improve their program and its impact? 8. Are the recommendations logical, actionable, and appropriate? 9. Have representative stakeholders and participants with diverse perspectives been involved in the interpretation process and formulating recommendations?
7. Report, disseminate, and use findings. Evaluators report on the findings, claims, and recommendations to the appropriate audiences, and engage or guide stakeholders in using the results to strengthen existing and future professional learning.	1. Will the evaluation have interim and/or final evaluation reports? 2. Who are the primary users of the evaluation report? 3. What components do the primary users want included in the evaluation report? 4. What format for reporting the results is most appropriate for the primary users of the evaluation report? 5. What other audiences are likely to want some version of the evaluation report? 6. What formats for reporting the results are appropriate for the other audiences? 7. Is the report sensitive to the human rights of participants (e.g., not including identifying information about individuals)? 8. How have other stakeholders and participants been involved in the reporting, disseminating, and use of the evaluation results? 9. Which groups are most likely to apply the results of this evaluation in their work? Have they been involved in learning about the evaluation results?
8. Evaluate the evaluation. As reflective practitioners, evaluators conduct a meta-evaluation of their efforts to strengthen their evaluation practice and inform future evaluations.	1. How will the effectiveness of the evaluation be assessed? 2. What questions will guide the evaluation of the evaluation? Consider credibility, validity, significance, resources, design, findings, and reporting. 3. What stakeholders will be involved in the evaluation of the evaluation? How will they be involved? 4. What key learnings about evaluation can be extracted from this evaluation that we want to apply to future evaluations? 5. What strengths are evident in the evaluator's practices, and what areas can be refined or modified?

Identifying KASABs

Delineating KASABs (knowledge, attitudes, skills, aspirations, and behaviors) is a way to define the outcomes of learning and the necessary changes required to achieve success with any initiative. In professional learning, KASAB defines the changes educators are expected to make to affect student success. Systemic change requires changes in KASABs for all key actors who contribute to, facilitate, lead, or are responsible for the change. For some initiatives, other actors such as parents and community members may also be expected to change.

This tool can be used in combination with the Mapping an Evaluation Step by Step tool. Fill out the desired outcomes for specific stakeholders. You will likely leave some cells blank.

Measurable outcomes	Students	Teachers	Coaches	Principals	Central office staff	Organization (policy, structures, systems, etc.)
Knowledge Conceptual understanding of information, theories, principles, and research.						
Attitudes Beliefs about the value of information or strategies.						
Skills The ability or capacity to use strategies and processes to apply knowledge.						
Aspirations Desires, or internal motivation, to engage in a practice.						
Behaviors Consistent application of knowledge and skills driven by attitudes and aspiration.						

Source: Killion, 2018.

Creating a logic model

Complete the table to create a logic model for your professional learning program, starting with listing the goal at the top. You may wish to use the sample logic model on p. 63 as a guide.

Professional learning program goal(s): _____

Inputs/ resources	Activities/ components	Outputs	Initial outcomes	Intermediate outcomes	Intended results
What resources, fiscal support, personnel, facilities, equipment, time, and technology do we need to accomplish the activities designed for this professional learning?	What is the sequence of actions we will take to achieve the outcomes of this professional learning?	What products, services, documents, or artifacts will we produce as we are engaged in the activities of this professional learning?	What are the initial changes in program participants we expect to see that, if present, will increase the likelihood of more substantial changes over time? (Usually changes in knowledge, skills, and attitudes.)	What are the intermediate changes in program participants we expect to see that, if present, will increase the likelihood of impact on students? (Usually changes in aspirations and behaviors.)	What are the expected changes in students? Does the degree of change vary over time?

Establishing an evaluation framework

To create an evaluation framework, start by listing the program goal. Then complete the table, using your answers to the questions in the Mapping an Evaluation Step by Step tool. You may wish to use the sample elements of an evaluation framework on p. 65 as a guide.

Professional learning program goal(s): _____

Measurable outcomes/ changes	Evaluation questions (formative and summative)	Data / evidence needed	Data source	Data collection method	Data analysis method	Timeline	Responsible person(s)