

BEST PRACTICES IN EVALUATING TEACHER PROFESSIONAL DEVELOPMENT

November 2015



In the following report, Hanover Research describes best practices for evaluating professional development for teachers and determining its impact on teacher effectiveness.



www.hanoverresearch.com

TABLE OF CONTENTS

Executive Summary and Key Findings	3
INTRODUCTION	3
KEY FINDINGS.....	4
Section I: Professional Development Trends	5
TRENDS IN PROFESSIONAL DEVELOPMENT	5
Program Effectiveness	5
Program Assessment	6
Section II: Professional Development Evaluation Models	8
LINEAR APPROACHES OF EVALUATION	8
DISTRIBUTED APPROACHES OF EVALUATION.....	10
IMPLEMENTATION CONSIDERATIONS	11
Usability of the Models.....	11
Identifying Change through Comparison.....	12
Data Collection.....	13
Section III: Sites of Evaluation.....	14
PARTICIPANT FEEDBACK.....	14
PARTICIPANT LEARNING.....	16
ORGANIZATIONAL CONTEXT	18
APPLICATION OF LEARNING	20
STUDENT OUTCOMES.....	20
Section IV: Organizational Profiles.....	23
WEST VIRGINIA DEPARTMENT OF EDUCATION.....	23
OHIO ADULT AND BASIC LITERACY EDUCATION.....	24
Appendix A	27
OHIO ABLE EVALUATION FRAMEWORK: LEVEL 1	27
OHIO ABLE EVALUATION FRAMEWORK: OPTIONAL LEVEL 2 COMPONENT.....	28
OHIO ABLE EVALUATION FRAMEWORK: LEVEL 3	28
Appendix B	30
GADSEN ELEMENTARY SCHOOL DISTRICT #32: ORGANIZATIONAL CONTEXT INDICATORS.....	30
Appendix C	32
WVDE SPECIAL EDUCATION TIS ADMINISTRATOR PRE SURVEY.....	32
WVDE SPECIAL EDUCATION TIS ADMINISTRATOR POST SURVEY	33

EXECUTIVE SUMMARY AND KEY FINDINGS

INTRODUCTION

According to Learning Forward, a national organization dedicated to fostering professional learning among teachers, “Professional learning that increases educator effectiveness and results for all students uses a variety of sources and types of student, educator, and system data to plan, assess, and evaluate professional learning.”¹ Using multiple data sources to analyze student, educator, and system performance helps create a more balanced and comprehensive portrait of the impact of professional learning programs, and this, in turn, positively contributes to program decision making.² Similarly, education experts from the Annenberg Institute for School Reform at Brown University recommend that educators approach issues related to school improvement by using a standardized inquiry protocol to collect data.³ For professional development activities, this inquiry framework helps provide a consistent approach to improvement that can be duplicated in different contexts.⁴

Within the context of this changing landscape, Hanover Research investigates best practices in evaluating teacher professional development to assist schools and districts in designing and evaluating their own teacher professional development processes. The report contains four sections:

- **Section I: Professional Development Trends** presents research on trends in professional development and addresses two potential areas for improvement.
- **Section II: Professional Development Evaluation Models** presents the structure and philosophy of three professional development evaluation models that have gained popularity in the education field.
- **Section III: Sites of Evaluation** offers an in-depth examination of common tools and approaches used to collect data from the evaluation areas identified in Section II.
- **Section IV: Organizational Profiles** provides profiles two educational organizations that have implemented evaluation frameworks based on the models discussed in this report.

¹ “Standards for Professional Learning.” Learning Forward. http://learningforward.org/standards#.U2p_4PIUg2

² “Data.” Learning Forward. http://learningforward.org/standards/data#.U49yx_ldVqU

³ Barnes, F. “Inquiry and Action: Making School Improvement Part of Daily Practice.” Annenberg Institute for School Reform at Brown University. 2004, p. 7. http://annenberginstitute.org/tools/guide/SIGuide_intro.pdf

⁴ Haslam, M.B. “Teacher Professional Development Evaluation Guide.” National Staff Development Council. 2010, pp. 62-63. <http://learningforward.org/docs/pdf/evaluationguide.pdf?sfvrsn=0>

KEY FINDINGS

- **Investment in professional development is growing; however, researchers highlight program effectiveness and assessment as two key development areas.** Estimates suggest that school districts in the United States spend between \$6 and \$18 billion on professional development. However, research suggests that it is difficult to isolate the role of professional development activities in improving teacher effectiveness.
- **The two most common approaches to professional development evaluation are based on different theories of teacher change.** Linear approaches posit a step-by-step change process wherein professional learning must precede changes in professional practice, while distributed approaches assume a more fluid dynamic that envisions learning and change in professional practice as an ongoing, back-and-forth process.
- **Regardless of the theory of teacher change that undergirds the evaluation framework, professional development evaluations generally aim to examine the following five areas:**
 - Participant Feedback
 - Participant Learning
 - Organizational Context
 - Application of Learning
 - Student Outcomes
- **Student outcomes represent one of the most critical and challenging sources of data for professional development evaluation.** Improved student outcomes are the ultimate goal of professional development, and as such provide a powerful testament to programs' success or failure. When selecting data sources, schools must be careful to ensure that the data capture the learning objectives covered in the training.
- **Assessment of participant learning helps to gauge the effectiveness of the training's pedagogical approach.** Possible data sources include pre- and post-training surveys or interviews, as well as more comprehensive tools such as scenario-based question prompts, quizzes, skills demonstrations, and portfolios. The tools used to assess the application of participant learning are similar, but tend to be administered over a longer time period to assess how quickly a teacher's practice may have changed.
- **An examination of organizational context is critical for understanding the drivers and impediments to teachers' implementing the skills acquired through professional development.** For instance, factors related to organizational context include the program's alignment with the district or school mission, its impact on organizational procedures, and the available resources to support teacher change, among others. Possible data sources for this evaluation area include school or district records and policy documents, meeting minutes, and surveys.

SECTION I: PROFESSIONAL DEVELOPMENT TRENDS

This section presents research on trends in professional development and addresses two potential areas for development.

TRENDS IN PROFESSIONAL DEVELOPMENT

Evidence from the United States Department of Education suggests that over the past decade, the proportion of discretionary funding devoted to professional development has increased.⁵ Estimates for professional development spending range from \$6 billion to \$18 billion.⁶ Additionally, a 2015 report from The New Teacher Project (TNTP) estimates that professional development activities required an average of 19 days of teacher time—roughly 10 percent of the school year.⁷

Research suggests that teachers both want and need quality professional development opportunities.⁸ Moreover, teachers who report having adequate time to collaborate with colleagues and participate in effective professional development experience higher job satisfaction rates than peers who lack these resources.⁹ However, some experts question whether these increased investments in professional development are justified, questioning the role of professional development as “the primary strategy for accelerating student learning.”¹⁰ These researchers argue that professional development programs carry an opportunity cost, since money that districts spend on professional development is unavailable to meet other needs.¹¹ The following subsections highlight two key development areas for professional development programs.

PROGRAM EFFECTIVENESS

Recent literature suggests that many professional development programs are ineffective in improving teacher performance. According to TNTP’s 2015 study of over 10,000 teachers,

⁵ “Findings From the 2013-14 Survey on the Use of Funds under Title II, Part A.” U.S. Department of Education. August 2014. p. 1. Available for download at: <http://www2.ed.gov/programs/teacherqual/resources.html>

⁶ [1] Crawford, Alvin, and Valerie Strauss. “The huge problem with professional development for teachers.” The Washington Post. September 6, 2014. <https://www.washingtonpost.com/blogs/answer-sheet/wp/2014/09/06/the-huge-problem-with-professional-development-for-teacher>

[2] “Teachers Know Best: Teachers’ Views of Professional Development.” Bill & Melinda Gates Foundation. December 2014. p. 3. <http://collegeready.gatesfoundation.org/wp-content/uploads/2015/04/Gates-PDMarketResearch-Dec5.pdf>

⁷ “The Mirage: Confronting the Hard Truth About Our Quest for Teacher Development.” The New Teacher Project. pp. 1, 4. http://tntp.org/assets/documents/TNTP-Mirage_2015.pdf.

⁸ “Teachers Know Best: Teachers’ Views of Professional Development,” Op. cit., p. 3.

⁹ Walker, Tim. “No More ‘Sit and Get’: Rebooting Teacher Professional Development.” National Education Association. April 29, 2013. <http://neatoday.org/2013/04/29/no-more-sit-and-get-rebooting-teacher-professional-development/>

¹⁰ “The Mirage: Confronting the Hard Truth About Our Quest for Teacher Development,” Op. cit., p. 2.

¹¹ “Findings From the 2013-14 Survey on the Use of Funds under Title II, Part A,” Op.cit.

despite an increase in professional development opportunities, nearly 70 percent of teachers observed received stable or declining evaluation ratings over 2-3 years. The organization notes that this lack of development was evident across all levels of teachers: for example, up to half of teachers with ten or more years of experience “were rated below ‘effective’ in core instructional practices, such as developing students’ critical thinking skills.”¹²

Additional researchers corroborate TNTP’s findings that many professional development opportunities are ineffective. An oft-cited 2009 study of teacher perceptions of professional development concludes that “fewer than 50% of teachers rated their professional development as useful.”¹³ The authors argue that the United States “is substantially behind other Organisation for Economic Co-operation and Development (OECD) nations in providing the kinds of powerful professional learning opportunities that are more likely to build their capacity and have significant impacts on student learning.”¹⁴

PROGRAM ASSESSMENT

Research suggests that in general, professional development evaluations do not directly account for teacher or student performance. A January 2014 report evaluating 643 studies on K-12 math professional development and found only five that met the rigorous What Works Clearinghouse (WWC) evidence standards. Of these, two concluded that the professional development program had a positive impact on math learning.¹⁵ A 2007 study by the Institute of Education Sciences (IES) concludes that nine out of over 1,300 studies met WWC standards for linking professional development to student achievement. The report concludes that “teachers who receive substantial professional development—an average of 49 hours in the nine studies— can boost their students’ achievement by about 21 percentile points.”¹⁶

Many researchers highlight difficulty isolating PD’s role in improving in teacher performance and student outcomes. As TNTP summarizes:

Even when teachers do improve, we were unable to link their growth to any particular development strategy. We looked at dozens of variables spanning the development activities teachers experienced, how much time they spent on them, what mindsets they brought to them and even where they worked. Yet we found no common threads that distinguished ‘improvers’ from other teachers. No type,

¹² Ibid., p. 2.

¹³ Darling-Hammond, Ruth, et al. “Professional Learning in the Learning Profession: A Status Report on Teacher Development in the U.S. and Abroad.” National Staff Development Council. February 2009. p. 59. <https://edpolicy.stanford.edu/sites/default/files/publications/professional-learning-learning-profession-status-report-teacher-development-us-and-abroad.pdf>

¹⁴ Ibid., p. 62.

¹⁵ Gersten, Russell, et al. “Summary of Research on the Effectiveness of Mathematics Professional Development Approaches.” Instructional Research Group, Institute of Education Sciences, U.S. Department of Education. January 2014. p. 1. http://ies.ed.gov/ncee/edlabs/regions/southeast/pdf/REL_2014010.pdf

¹⁶ Yoon, Kwang Suk, et al. “Reviewing the evidence on how teacher professional development affects student achievement.” Institute of Education Studies. October 2007. P. iii. http://ies.ed.gov/ncee/edlabs/regions/southwest/pdf/rel_2007033.pdf

amount or combination of development activities appears more likely than any other to help teachers improve substantially, including the ‘job-embedded,’ ‘differentiated’ variety that we and many others believed to be the most promising.¹⁷

Recent academic studies note that most assessments of teacher professional development rely on teacher satisfaction and perceptions of quality rather than objective measures of teacher learning, classroom use of new practices, or improvements in student performance.¹⁸ In their 2011 study of professional development program effects on high school biology teachers, Diane Ebert-May, et al. observe that “there is still an absence of objective measures for evaluating teaching practices of faculty following professional development.”¹⁹ The lack of these objective assessments has led several researchers to describe evaluations as “the weakest link in the PD chain.”²⁰

Despite these concerns, research continues to confirm the importance of professional development opportunities for teachers. According to researchers Gersten, et al., schools seeking to implement effective professional development “must use their best judgment until more causal evidence becomes available.”²¹ Therefore, the remainder of this report addresses best practices and strategies for evaluating the effectiveness of PD.

¹⁷ “The Mirage: Confronting the Hard Truth About Our Quest for Teacher Development,” Op. cit., p. 2.

¹⁸ [1] King, Fiona. “Evaluating the impact of teacher professional development: an evidence-based framework.” *Professional Development in Education*. 40.1. 2014. Available for download at: <http://www.tandfonline.com/doi/full/10.1080/19415257.2013.823099>

[2] Ebert-May, Diane, et al. “What We Say Is Not What We Do: Effective Evaluation of Faculty Professional Development Programs.” *BioScience*. 61.7. (2011). p. 550.
<http://bioscience.oxfordjournals.org/content/61/7/550.full.pdf+html>

¹⁹ Ebert-May et al., Op. cit., p. 550.

²⁰ King, Op. cit.

²¹ Gersten et al., Op. cit., p. 1.

SECTION II: PROFESSIONAL DEVELOPMENT EVALUATION MODELS

In the past decade, three professional development evaluation models have gained popularity in the education field: the Kirkpatrick, Guskey, and Clarke-Hollingsworth models. Each of these models differs based on its conceptualization of “teacher change.” Generally, teacher change is founded on the idea that teachers implement their professional practice as learners in a broader learning community of educators. As a result, professional development has the ability to change teacher professional practice through contributing to and fostering learning.²²

Each evaluation model discussed in this report has a unique causal chain that enumerates how teacher learning, student impact, and professional development activities influence each other.²³ The Kirkpatrick and Guskey models are both *linear* approaches to teacher change, wherein one stage in teacher change leads directly to another. In contrast, the Clarke-Hollingsworth model is a *distributed* approach, wherein each component of teacher change is linked to others and can be initiated from multiple points.²⁴ This section explores the structures of each model and explains how the different conceptualizations of teacher change influence each approach to evaluation.

LINEAR APPROACHES OF EVALUATION

Developed by Dr. Donald Kirkpatrick, Professor Emeritus at the University of Wisconsin and former President of the American Society for Training and Development (ASTD), the Kirkpatrick Training Evaluation Model has been adapted to a wide range of educational and industry contexts since the late 1950s.²⁵ The Kirkpatrick model includes four levels: instead of measuring only one outcome at the end of the program, it focuses on measuring four types of tiered training outcomes. These outcomes include participant reactions, participant learning, participant behavior, and student results (Figure 2.1).²⁶

²² Clark, D. and Hollingsworth, H. “Elaborating a model of teacher professional growth.” *Teaching and Teacher Education*, 18, 2002, p. 948.

<https://xa.yimg.com/kq/groups/73868647/1021761450/name/clark+and+Hollingsworth.pdf>

²³ *Ibid.*, p. 949

²⁴ Warren, E. “Early Childhood Teachers’ Professional Learning in Early Algebraic Thinking: A Model that Supports New Knowledge and Pedagogy.” *Mathematics Teacher Education and Development*, 10, 2008/2009. pp. 30 and 31. <http://eric.ed.gov/?id=EJ863711>

²⁵ [1] “About US: Donald L. Kirkpatrick.” Kirkpatrick Partners.

<http://www.kirkpatrickpartners.com/AboutUs/DonKirkpatrick/tabid/223/Default.aspx>

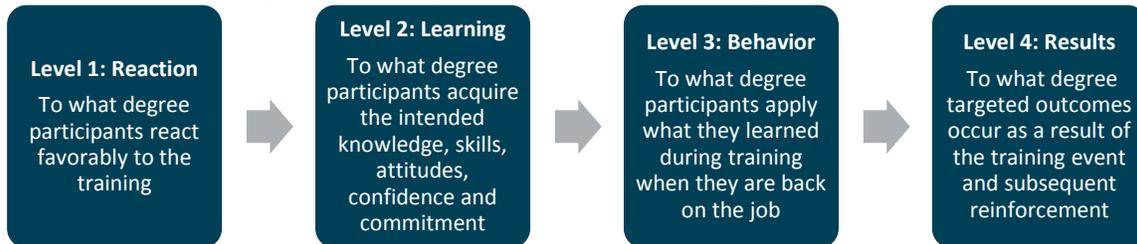
[2] “Kirkpatrick’s Four-Level Training Evaluation Model.” Mind Tools.

<http://www.mindtools.com/pages/article/kirkpatrick.htm>

²⁶ Preskill, H. and Russ-Eft D. “Building Evaluation Capacity.” Sage Publications, Thousand Oaks, CA. 2004, p. 101.

http://www.sagepub.com/upm-data/5068_Preskill_Chapter_5.pdf

Figure 2.1: Kirkpatrick Four-Level Training Evaluation Model

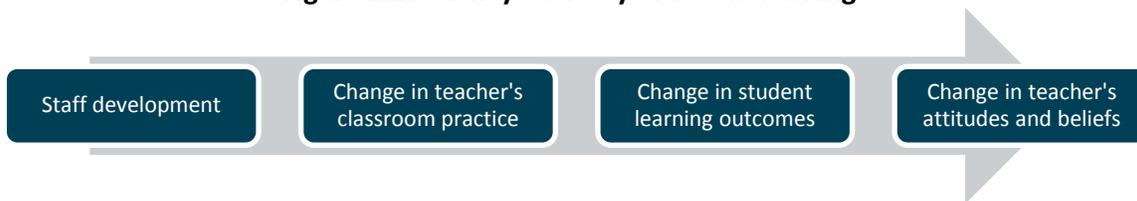


Source: Kirkpatrick Partners²⁷

The Kirkpatrick model operates on the implicit assumption that the causal chain of teacher learning is sparked by a receptive reaction to the information the teacher acquires in class. That is, professional development causes teachers to modify their knowledge and beliefs, which in turn causes them to change their behavior and classroom practices. At the end of the causal chain, if the previous steps are followed, students demonstrate better outcomes because of improved instructional and professional practices on the part of their teachers.²⁸ Accordingly, the Kirkpatrick evaluation model is intended as an accumulative process that builds on the data collected at each previous level, and aims to provide a more detailed layer of assessment at each successive level.²⁹

The Guskey model is based on the Kirkpatrick model, but was developed explicitly for an educational setting and follows a somewhat different conceptualization of the causal chain of teacher change. The Guskey model was developed by Dr. Thomas Guskey, a Professor of Education Psychology at the University of Kentucky’s College of Education.³⁰ In Guskey’s conceptualization of causal change, shifts in teacher attitude and knowledge do not occur solely because of the information acquired in a training session. Rather, “teachers change their beliefs and attitudes through changing their practice and reflecting on the results” (Figure 2.2).³¹

Figure 2.2: Guskey’s Theory of Teacher Change



Source: Clarke and Hollingsworth³²

²⁷ Taken verbatim from “The Kirkpatrick Model.” Kirkpatrick Partners. <http://www.kirkpatrickpartners.com/OurPhilosophy/TheKirkpatrickModel/tabid/302/Default.aspx>

²⁸ Clarke and Hollingsworth, Op. Cit., p. 949.

²⁹ “Evaluating Training Programs: Kirkpatrick’s Four Levels.” Washington State Employment Security Department. http://www.wa.gov/esd/training/toolbox/tg_kirkpatrick.htm

³⁰ “Educational, School, and Counseling Psychology: Thomas Guskey.” University of Kentucky College of Education. <https://2b.education.uky.edu/edp/educational-psychology-overview/educational-psychology-faculty/>

³¹ Guskey (1986) from Clarke and Hollingsworth, Op. Cit., p. 50.

³² Ibid.

Guskey suggests that once teachers see the power of a new teaching method, they are more likely to believe that the method is effective and continue to apply it, which creates a positive self-perpetuating cycle. Therefore, creating space for teachers to effectively implement new practices in their classrooms and directly evaluate student learning is critical.³³

Dr. Guskey further asserts that Kirkpatrick’s model did not fully illustrate the process by which professional development contributes to improved student outcomes because it did not account for the critical role that resources and the school or district environment play in the professional development of teachers.³⁴ Thus, Guskey created one additional level of evaluation to the Kirkpatrick framework, called “Organizational Change and Support,” which aims to identify the means by which professional development lessons are embedded in the structural organization of an entity. This additional level allows evaluators to examine a teacher’s access to resources and institutional support, both of which assist in the application of new ideas.³⁵ Figure 2.3 provides a diagram of Guskey’s modified approach to professional development evaluation, which, despite having a different causal chain than Kirkpatrick, follows a similar evaluation process.

Figure 2.3: Guskey’s Five Levels of Professional Development Evaluation



Source: Guskey (2000)³⁶

DISTRIBUTED APPROACHES OF EVALUATION

As described above, the Guskey model provides an alternative to the idea that it is necessary to change teacher beliefs and attitudes *prior* to changing classroom practices in order to achieve improved student outcomes. The Clarke-Hollingsworth conceptualization of teacher change builds on Guskey’s model by eliminating the hierarchical levels and structuring the evaluation framework according to domains. This model assumes that the process of teacher change can be initiated from changes occurring in any domain, instead of having to follow a linear path.³⁷ According to this model, instruction is influenced by the Domain of Practice, or instructional behavior; the Personal Domain, or the beliefs that prompt behavior; the Domain of Consequence, or the beliefs about what will happen as a

³³ Ibid., 31

³⁴ Kreider, H. and Bouffard, S. “A Conversation with Thomas R. Guskey.” *Harvard Evaluation Exchange*. XI:4, Winter 2005/2006. <http://www.hfrp.org/evaluation/the-evaluation-exchange/issue-archive/professional-development/a-conversation-with-thomas-r.-guskey>

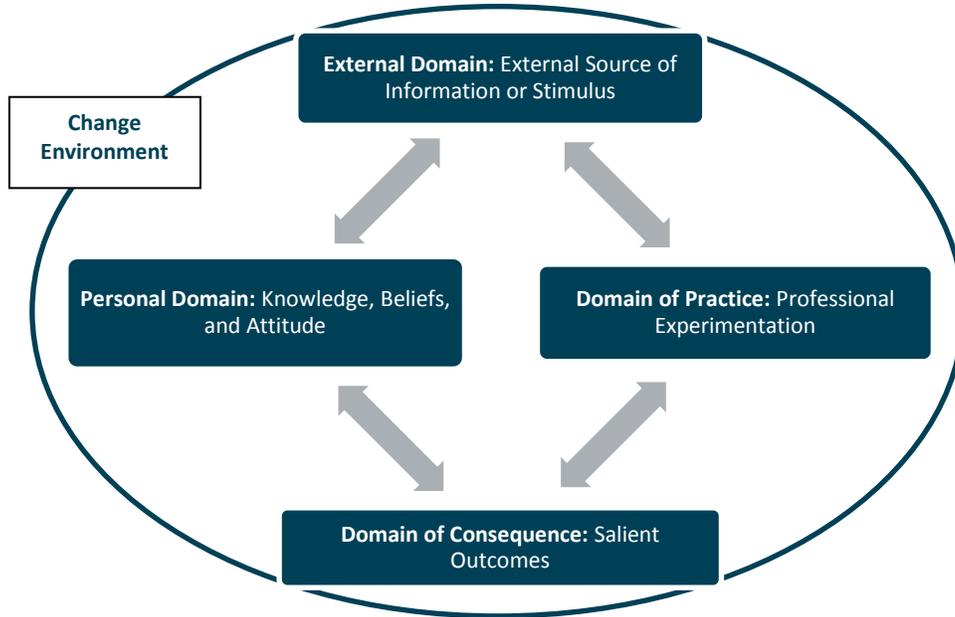
³⁵ Ibid.

³⁶ Guskey (2000) from Ibid.

³⁷ Clarke and Hollingsworth, Op. Cit., p. 949.

result of those actions; and the External Domain, which is the site of new information and stimuli (e.g., professional development activities) that can influence practice (Figure 2.4).³⁸

Figure 2.4: Clarke-Hollingsworth Model



Source: Clarke and Hollingsworth³⁹

In the Clarke-Hollingsworth model, change occurs through “mediating processes of reflection and enactment.” That is, change in one domain triggers a teacher to reflect critically on professional practice and make changes to other domains. However, the process of a teacher engaging in reflection and change also depends on the overall Change Environment.⁴⁰ Change Environment refers to the supports or impediments to change that a teacher faces, and is a concept that is similar to Guskey’s Organization Support and Change level. Although it was not explicitly created as an evaluation model, the Clarke-Hollingsworth model of teacher growth is designed as an analytical tool that facilitates the assessment of professional development activities.⁴¹

IMPLEMENTATION CONSIDERATIONS

USABILITY OF THE MODELS

The models discussed in the report require the examination of a range of data and provide frameworks for both summative and formative evaluations. Summative evaluations are generally conducted at the end of a program to determine its overall effectiveness.

³⁸ McDonough, A. and Clarke, B. “Professional development as a Catalyst for Changes in Beliefs and Practice: Perspectives from the Early Numeracy Research Project.” p. 521. www.merga.net.au/documents/RP582005.pdf

³⁹ Adapted from Ibid., p. 951

⁴⁰ Clarke and Hollingsworth, Op. Cit., p. 951

⁴¹ Ibid., p. 958

Conversely, formative evaluations are conducted during a program to gain insight into how the program is being implemented and to gather information that can be used to help modify or replicate the program in the future.⁴² In particular, formative evaluations rely on an understanding of how participant learning and professional practice change over the course of professional development. In contrast, summative evaluations are more likely to focus on student outcome data in order to understand the overall impact of professional development activities.⁴³

Whether an evaluation’s purpose is formative or summative will likely affect the design of and the resources invested in the evaluation process. For instance, in an analysis of a campus that implemented the Guskey evaluation framework and was awarded with the U.S. Department of Education Model Professional Development Award, researchers found evidence of evaluation for all five levels of the model.⁴⁴ However, while evaluation of Participants’ Reactions (Level 1) and Participants’ Learning (Level 2) could be achieved simply by using standardized forms, evaluation of Organization Support and Change, Use of New Knowledge and Skills, and Student Learning Outcomes (Levels 3-5) required additional investments of funds and time from school leadership. This is likely because the evaluation of Levels 3-5 relied more heavily on school artifacts and interviews with teachers and staff.⁴⁵

IDENTIFYING CHANGE THROUGH COMPARISON

When distilled to their basic components, the models described in this report are intended to determine if professional development led teachers to change their professional practice, and if so, whether that change affected student outcomes. In the evaluation literature, comparisons are the primary method used to measure change. In *Assessing Impact*, Joellen Killion provides a summary table of the most common kinds of comparisons used in impact evaluations (Figure 2.5).

Figure 2.5: Common Comparisons for Impact Evaluations

TITLE	DESCRIPTION
Individual Comparison	Compare scores from same individual pre- and post-intervention
Cohort Group Comparison	Compare scores from same group pre- and post-intervention
Panel Group Comparison	Compare post-intervention scores of same group from two different years (e.g., first graders in one year vs. first graders the next year)
Selected Comparison (Control Trial) Group	Compare pre- and post-intervention scores of an intervention and a control group

Source: Joellen Killion⁴⁶

⁴² Killion, J. “Assessing Impact, 2nd Ed.” Corwin Press, A Sage Company: Thousand Oaks, CA. 2008. pp. 13 and 15.

⁴³ Ibid.

⁴⁴ McMahan, M. “A Case Study of a School Organization as it Relates to Staff Development Evaluation.” Texas A&M University Thesis, May 2000. p. v.

⁴⁵ Ibid., p. 3.

⁴⁶ Killion, Op. Cit., p. 75.

Each of these comparison types has certain advantages and disadvantages. For instance, individual comparisons are useful in providing information about changes, but they may not allow for conclusions about attribution. Additionally, cohort group comparisons, panel group comparisons, and selected comparisons provide information about the impact of the program, but “may not account of differences present in the groups before the staff development intervention or interventions that may have been used in the comparison group during the period of interest [emphasis in original].”⁴⁷

DATA COLLECTION

The process of data collection can be simplified if techniques and instruments are piloted with a small group and if data collectors are well-trained on the associated expectations and methods.⁴⁸ As a first step, evaluators should examine the professional development program’s logic model or theory of change to determine what kind of data to collect.⁴⁹ Additionally, regardless of the source of data, Killion further recommends that evaluators ensure that they understand the answers to each of the following five questions when collecting data:

1. Are the data being collected those that were planned?
2. What problems are occurring in the data-collection process, and how can they be resolved?
3. What other data might need to be collected?
4. How do I manage data during this time?
5. How can I ensure accuracy and precision in the data-collection process?⁵⁰

In addition to piloting the data collection instruments and processes, Killion also notes that evaluators should ensure that there is clarity regarding the management of data collection, including a plan and master schedule that “delineates who will collect the data, where the data will be collected, when the data will be collected, and how the data will be collected.”⁵¹

⁴⁷ Ibid.

⁴⁸ Ibid., pp. 97-98

⁴⁹ Haslam, Op. Cit., p. 24.

⁵⁰ Killion, Op. Cit., p. 97

⁵¹ Ibid., p. 98.

SECTION III: SITES OF EVALUATION

Linear and distributed approaches to professional development evaluation are based on different causal chains of teacher change; however, these approaches to evaluation examine many of the same core areas. This section reviews the strategies and tools used to collect data for each of the evaluation areas associated with the Guskey, Kirkpatrick, and Clarke-Hollingsworth models (Figure 3.1).

Figure 3.1: Evaluation Model Crosswalk

EVALUATION AREA	GUSKEY MODEL	KIRKPATRICK MODEL	CLARKE-HOLLINGSWORTH MODEL
Participant Feedback	Level 1: Participants’ Reactions	Level 1: Reaction	N/A
Participant Learning	Level 2: Participants’ Learning	Level 2: Learning	Personal Domain and External Domain
Organizational Context	Level 3: Organization Support and Change	N/A	Change Environment
Application of Learning	Level 4: Participants Use of New Knowledge and Skills	Level 3: Behavior	Domain of Practice
Student Outcomes	Level 5: Student Learning Outcomes	Level 4: Results	Domain of Consequences

Source: Kirkpatrick Partners and Connecting Communities⁵²

PARTICIPANT FEEDBACK

The Kirkpatrick and Guskey models both assert that participants should provide feedback on the training.⁵³ In particular, the Kirkpatrick model takes an in-depth approach to this issue in that it also seeks to understand how trainees *experienced* the training overall.⁵⁴ Kirkpatrick dictates that there are three main areas at this level that evaluators should examine:

- **Satisfaction:** Obtain insight into how well participants liked the training. This element is also present in the Guskey model.
- **Engagement:** Measure the extent to which participants were actively involved in and contributed to the training.
- **Relevance:** Examine the types of future opportunities participants believe they will have in the course of their work to apply training lessons.⁵⁵

⁵² [1] “Guskey’s Five Critical Levels of Professional Development Evaluation.” Connecting Communities.

<http://connectingcanycommunities.wikispaces.com/file/view/Guskey+5+levels.pdf>

[2] “The Kirkpatrick Model,” Op. Cit.

⁵³ [1] Guskey, T. “Does It Make A Difference? Evaluating Professional Development.” *Educational Leadership*, 59:6 2002. <http://www.ascd.org/publications/educational-leadership/mar02/vol59/num06/Does-It-Make-a-Difference%C2%A2-Evaluating-Professional-Development.aspx>

[2] “Kirkpatrick’s Four-Level Training Evaluation Model,” Op. Cit.

⁵⁴ Ibid.

⁵⁵ “The New World Kirkpatrick Model.” Kirkpatrick Partners.

<http://www.kirkpatrickpartners.com/OurPhilosophy/TheNewWorldKirkpatrickModel/tabid/303/Default.aspx>

Kirkpatrick and Guskey both rely heavily on the use of questionnaires and surveys to collect data at this level and encourage the use of mixed methods in analysis.⁵⁶ Kirkpatrick argues that feedback forms should allow evaluators to quantify responses (e.g., “60 percent of respondents think that...”), while also providing space for qualitative feedback through written comments and suggestions.⁵⁷ Similarly, Guskey recommends using questionnaires with a combination of Likert-type rating scales and open-ended questions.⁵⁸

Figure 3.2 provides a comparison of the examples of questions that both models recommend asking participants at this level. Notably, while the rationale for inclusion of questions directly related to training content is self-evident, the Guskey model also encourages asking questions focused on physical comfort. These questions allow evaluators to assess how well a training session prepared participants for learning.⁵⁹ An example of how one educational program translated these sample questions into a questionnaire can be found in Appendix A and is discussed in the Ohio ABLE case study in Section IV. Kirkpatrick also asserts that in addition to using inquiry-based tools, evaluators can indirectly assess Engagement and Trainee Satisfaction by observing participant body language during training sessions.⁶⁰

Figure 3.2: Examples of Questions to Gauge Participant Reaction

GUSKEY ⁶¹	KIRKPATRICK ⁶²
<ul style="list-style-type: none"> ■ Did participants like it? ■ Was their time well spent? ■ Did the material make sense? ■ Were the activities well planned and meaningful? ■ Was the leader knowledgeable and helpful? ■ Did the participants find the information useful? ■ Questions related to physical comfort: <ul style="list-style-type: none"> ○ Were the refreshments good? ○ Was the room at the right temperature? ○ Were the chairs comfortable? 	<ul style="list-style-type: none"> ■ Did the trainees feel that the training was worth their time? ■ Did they think that it was successful? ■ What were the biggest strengths of the training? ■ What were the biggest weaknesses of the training? ■ Did they like the venue and presentation style? ■ Did the training session accommodate their personal learning styles?

⁵⁶ [1] Guskey (2002), Op. Cit.

[2] “Kirkpatrick’s Four-Level Training Evaluation Model,” Op. Cit.

⁵⁷ “Evaluating Training Programs: Kirkpatrick’s 4 Levels,” Op. Cit.

⁵⁸ Guskey (2002), Op. Cit.

⁵⁹ Ibid.

⁶⁰ Ibid.

⁶¹ [1] Ibid.

[2] Guskey (2000) from “A Conversation with Thomas R. Guskey,,” Op. Cit.

⁶² Taken verbatim from “Kirkpatrick’s Four-Level Training Evaluation Model,” Op. Cit.

PARTICIPANT LEARNING

Both linear and distributed approaches encourage evaluators to examine the knowledge and skills that participants have acquired, whether this acquisition occurred directly through training or indirectly through reflection and changes in a different domain.⁶³ For instance, Clarke and Hollingsworth emphasize the importance of participants’ beliefs and attitudes, while Kirkpatrick explicitly encourages evaluators to measure attitude, confidence, and commitment, which, alongside knowledge and skills, make up five distinct areas of learning.⁶⁴ Figure 3.3 provides a first-person definition of the significance of each area as described by the Kirkpatrick model.⁶⁵

Figure 3.3: Kirkpatrick Learning Areas

LEARNING AREA	DESCRIPTION
Knowledge	“I know it.”
Skill	“I can do it right now.”
Attitude	<i>“I believe this will be worthwhile to do on the job.”</i>
Confidence	“I think I can do it on the job.”
Commitment	<i>“I intend to do it on the job.”</i>

Source: Kirkpatrick Partners⁶⁶

The Guskey model has a single overarching question that guides this level: “Did participants acquire the intended knowledge and skills?”⁶⁷ Guskey dictates that assessments require active demonstration on the part of participants.⁶⁸ Therefore, he suggests using tools aligned with training content objectives that require participants to provide evidence of what they have learned, such as:

- Pencil-and-paper quizzes
- Skills demonstrations
- Written and oral reflections
- Portfolios assembled over the course of the training⁶⁹

Additionally, a 2009 article published on the evaluation of the professional development program Transformative Teaching in Early Years Mathematics (TTEYM) illustrates that many of the five learning areas can be assessed through written and oral reflection, and that interviews can provide both qualitative and quantitative data. TTEYM is grounded in the Clarke and Hollingsworth and the Guskey theories of professional development and teacher

⁶³ [1] Ibid.

[2] Guskey (2002), Op. Cit.

[3] Clarke and Hollingsworth, Op. Cit., p. 951.

⁶⁴ “The New World Kirkpatrick Model,” Op. Cit.

⁶⁵ Ibid.

⁶⁶ Ibid.

⁶⁷ Guskey (2000), taken from “A Conversation with Thomas R. Guskey,,” Op. Cit.

⁶⁸ Guskey (2002), Op. Cit.

⁶⁹ Guskey (2000), taken from “A Conversation with Thomas R. Guskey,,” Op. Cit.

change.⁷⁰ The six-month program consisted of two cycles, each of which used an expert to guide teachers through different facets of lesson plan development and implementation.⁷¹ The authors of the study used the following tools to assess the learning of teachers and students:

- Interviews conducted with teachers at the end of the six-month program and then again 18 months later
- Field notes written during classroom observations
- Videotape recordings of lessons and professional development activities
- Interviews conducted with a sample of students at the end of each cycle⁷²

The first set of teacher interviews were open-ended and conducted by a neutral third party to encourage respondents to share their opinions on the new model. Transcripts of the interviews were coded and analyzed to determine the concepts mentioned with the highest frequency.⁷³ The second set of interviews drew from the three categories with the highest mention frequencies (i.e., mathematical knowledge, mathematical thinking, and personal confidence), and engaged in deeper open-ended questioning about how teachers' beliefs, attitudes, and skills had changed in each area.⁷⁴ Conducting the interview sets at different times allowed for longitudinal comparison of teachers' beliefs and knowledge about mathematics, thus providing more precise insights into the type of learning that the program had sparked.

Scenarios are another tool that can be useful as a measure of learning.⁷⁵ Scenario-based question prompts simulate real-life situations through detailed descriptions of particular challenges or situations. They are a popular tool for educator professional development because they require participants to practice new skills and help them to gain confidence in applying new knowledge and skills in the workplace.⁷⁶ There are primarily two types of scenarios. In a *clinical* scenario, the participant responds to a set of questions that have predetermined and measurable answers, similar to standardized assessments. Clinical scenarios are often used in medical education. In contrast, *situational* scenarios require the learner to respond to an open-ended question about how they would use a skill in the workplace.⁷⁷ Because there are numerous ways of effectively applying new professional development knowledge and skills in the classroom, situational scenarios are often better suited to the educational context.

⁷⁰ Warren, Op. Cit., p. 1.

⁷¹ Ibid., p. 37.

⁷² Ibid.

⁷³ Ibid.

⁷⁴ Ibid., p. 42.

⁷⁵ Ibid., p. 165.

⁷⁶ Nichols, A. "Pre- and Post-Scenarios: Assessing Learning and Behavior Outcomes in Training Settings." *The Canadian Journal of Program Evaluation*, 19:2, 2004. p. 167.

⁷⁷ Ibid., p. 167.

In the area of learning, the primary method of scenario analysis is what Joellen Killion refers to as a cohort comparison (see Figure 1.5). Scenarios are given to participants before training to establish a baseline of ability, and then again after training to determine the extent to which participants are able to apply their newly learned knowledge and skills to relevant situations.⁷⁸ Additionally, analysis of change in clinical scenarios can be conducted by using statistical tests, while situational scenario comparisons can be done by tracking the appearance of themes or key concepts.⁷⁹

ORGANIZATIONAL CONTEXT

In the Organization Support and Change level in Guskey's model and in the Change Environment of Clarke and Hollingsworth's model, evaluators examine the broader context in which professional development and professional practice take place. Guskey places particular focus on a participant's work environment because organizational policies and culture can support or undermine efforts to implement new skills.⁸⁰ For example, if funding is unexpectedly reduced for classroom aides, than full-time teachers may experience greater demands on their time and be less able to attend professional development sessions or experiment with new instructional methods.⁸¹ While Kirkpatrick does not have an explicit evaluation area dedicated to organizational context, in Level 4: Behavior, he encourages evaluators to bear in mind the context of a trainee's overall environment because behavioral change only happens when "required drivers," such as systems that reinforce and reward certain behaviors, are present.⁸²

Data collection related to organizational context requires evaluators to draw on a broad range of sources, such as:

- District or school records
- Minutes from follow-up meetings
- Questionnaires and interviews provided to both participants and school administrators containing questions such as:
 - Did professional development promote changes that were aligned with the mission of the school and district?
 - Did it affect organizational climate and procedures?
 - Was implementation advocated, facilitated, and supported?
 - Was the support public and overt?
 - Were changes at the individual level encouraged and supported at all organizational levels?
 - Were sufficient resources made available, including time for sharing and reflection?

⁷⁸ Ibid., p. 171.

⁷⁹ Ibid., pp. 171 and 173.

⁸⁰ Guskey (2002), Op. Cit.

⁸¹ Haslam, Op. cit., p. 26

⁸² "The New World Kirkpatrick Model," Op. Cit.

- Were successes recognized and shared?
- Were problems addressed quickly and efficiently?⁸³

Guskey's commitment to examining Organization Support and Change also aligns with Learning Forward's recommendation to compare a training's original plan to how it was actually implemented in order to identify important impediments to professional development.⁸⁴ In particular, Learning Forward recommends that evaluators collect data on:

- The availability of supplies and equipment to implement new professional practices.
- How contextual factors such as changes in leadership, school or district priorities, resources, teacher assignments, or student body demographics influenced implementation.
- The extent to which all players identified in professional development, including participants, presenters, facilitators, administrative staff, and leadership, carried out their responsibilities for professional development.⁸⁵

The Gadsen Elementary School District provides an example of how the multiple characteristics that comprise organizational context can be measured using a unified assessment framework. In this example, evaluators assess the role that leadership plays in shaping organizational context by placing school leadership on a performance continuum in the following six critical areas:

- Create atmosphere or context for change
- Develop and communicate a shared vision
- Plan and provide resources
- Invest in professional development
- Check for progress
- Provide assistance⁸⁶

For example, in the category "Create atmosphere or context for change," leadership that simply creates time for collaborative work would be placed at the lower end of the scale, while leadership that creates time for collaborative work while also helping to foster staff skills in the areas of collaboration, modes of conversation, conflict management, and decision making would be placed at the higher end of the spectrum.⁸⁷ A copy of the full assessment spreadsheet and the indicators used at Gadsen Elementary School District can be found in Appendix B.

⁸³ [1] Ibid.

[2] Guskey (2000), Op. Cit.

⁸⁴ Haslam, Op. Cit. p. 24.

⁸⁵ Ibid., p. 25

⁸⁶ Killion, Op. Cit., pp. 187.

⁸⁷ Ibid., pp. 187-189.

APPLICATION OF LEARNING

The Kirkpatrick model refers to this level as Behavior, while in the Guskey model it is known as Participants' Use of New Knowledge and Skills and in Clarke and Hollingsworth's model it is considered the Domain of Practice or professional experimentation. In each model, this area of evaluation is designed to assess how instructors apply what they have learned to their professional practice. Similar to the learning evaluation area, in the Guskey model there is one overarching question related to the application of learning: "Did participants effectively apply the new knowledge and skills?" Guskey further notes that questionnaires, structured interviews with participants and supervisors, written reflections, and examination of portfolios are all acceptable ways to assess this level.⁸⁸ Kirkpatrick provides specific examples of the questions that evaluators can ask to assess this evaluation area:

- Did the trainees put any of their learning to use?
- Are trainees able to teach their new knowledge, skills, or attitudes to other people?
- Are trainees aware that they have changed their behavior?⁸⁹

Guskey and Kirkpatrick both encourage evaluators to collect data for a prolonged period after the end of a training in order to allow time for teachers to modify their professional practice, and to observe how and in what increments they do so. In the Kirkpatrick model, evaluators are expected to engage in data collection for three to six months, while in the Guskey model the amount of time allotted depends on evaluator preferences.⁹⁰

The West Virginia Department of Education provides an additional example of how the application of learning can be measured through the provision of pre-tests and post-tests that compare changes in teacher responses concerning the usage of certain strategies in class. Full samples of these tests can be found in Appendix C. Ohio ABLE also gauges the application of learning through a single, reflective questionnaire administered after the program. This sample questionnaire can be found in Appendix A. In-depth descriptions of both the West Virginia and Ohio programs are presented in Section IV of this report.

STUDENT OUTCOMES

There is a growing body of literature that links teacher learning and professional development with improved student achievement. However, testing the precise nature of the relationship between professional development and student outcomes is challenging because there are typically confounding factors that may create confusion or uncertainty over perceived linkages.⁹¹ The prevalence of confounding factors makes examination of the

⁸⁸ Guskey (2002), Op. Cit.

⁸⁹ Taken verbatim from "Kirkpatrick's Four-Level Training Evaluation Model," Op. Cit.

⁹⁰ [1] Guskey (2000), Op. Cit.

[2] Kirkpatrick, D. "Implementing Kirkpatrick's Four Levels TU101." ASCD, 2007. p. 12.

<http://astd2007.astd.org/PDFs/Handouts%20for%20Web/Handouts%20Secured%20for%20Web%205-15%20thru%205-16/TU101.pdf>

⁹¹ Reitzug Ulrich C. "Professional Development." In A. Molnar (ed.), "School Reform Proposals: The Research Evidence." National Education Policy Center, January 1, 2002. pp. 5 and 6.

<http://nepc.colorado.edu/files/Chapter12-Reitzug-Final.pdf>

four previous content areas that undergird theories of teacher change especially important. After all, simply measuring student achievement before and after a program does not enumerate the *processes* by which professional development makes an impact on student outcomes.⁹²

In addition, Guskey points out that improved academic achievement are not the only positive student outcomes that can accrue from effective professional development. Specifically, he recommends using the following questions, which extend beyond the realm of academic achievement per se, to guide data collection in this area:

- What was the impact on students?
- Did it affect student performance or achievement?
- Did it influence students' physical or emotional well-being?
- Are students more confident as learners?
- Is student attendance improving?
- Are dropouts decreasing?⁹³

In the Kirkpatrick model, evaluators are asked to answer similar questions, but are encouraged to use specific “leading indicators.” These leading indicators are short-term observations and measurements that show whether performed behaviors learned from training are actually creating the desired impact.⁹⁴ The selection of indicators and measurement practices at this level is highly dependent on the desired outcomes for the training.⁹⁵

Additionally, Guskey notes, somewhat counterintuitively, that evaluations should also include indicators that are unrelated to the content of the training. For example, after participating in writing workshops, educators might develop new lesson plans to improve students’ writing. Writing scores might improve, but at the same time math scores might decline because the teachers’ new skills require additional instructional time to implement. An assessment of student outcomes that only looks at writing and reading scores might miss this critical but unintended consequence of the training. Guskey recommends examining indicators at the individual, class, school, and even district levels, depending on the nature of the training.⁹⁶

Frequently, evaluators select data sources that are easily accessible such as standardized test scores or student grades. However, this practice may not provide measures of student achievement that are sensitive enough to pick up the unique impact of professional

⁹² InPraxis Group Inc. “Effective professional Development: What the Research Says.” Alberta Education, 2006. p. 36. <http://www.assembly.ab.ca/lao/library/egovdocs/2006/aled/158121.pdf>

⁹³ Guskey, (2000), Op. Cit.

⁹⁴ “Kirkpatrick’s Four-Level Training Evaluation Model,” Op. Cit.

⁹⁵ Ibid.

⁹⁶ Ibid.

development. For example, standardized tests are generally highly reliable psychometrically, but they are designed to assess broad content retention and not specific curricular indicators, often rendering them ill-suited for professional development evaluation purposes. The measure of student achievement that evaluators use should be aligned not only with the broader curriculum as well as classroom instructional practices and assessments, but also with the content covered in the professional development.⁹⁷

Common student outcome data sources include:

- Samples of student work
- Student scores on local benchmark assessments
- Student scores on state assessments
- Student and school records
- Structured interviews with students and parents⁹⁸

Furthermore, one method of using samples of student work to assess impact on student achievement is to include them in teacher portfolios. Analysis of changes in student samples over time can indicate whether they are consistent with observed changes in participant learning and the application of learning.⁹⁹

⁹⁷ Killion, Op. Cit., p. 29.

⁹⁸ [1] Haslam, Op. Cit., p. 31.

[2] Guskey, (2000), Op. Cit.

⁹⁹ For an example school portfolio guidance document, see: Killion, Op. cit., p. 178.

SECTION IV: ORGANIZATIONAL PROFILES

This section profiles two educational organizations that have implemented evaluation frameworks based on the models discussed in Section II and Section III of this report.

WEST VIRGINIA DEPARTMENT OF EDUCATION

The West Virginia Department of Education (WVDE) Technology Integration Specialist (TIS) program provides professional development activities related to 21st Century Technology Tools to create school-based technology specialists.¹⁰⁰ In 2006, the program expanded to include special education practitioners by providing 320 hours of specialized online and face-to-face professional development sessions.¹⁰¹ WVDE developed an evaluation framework based on the Guskey model to assess the effectiveness of the TIS Special Education program.¹⁰² WVDE’s adaptation of the Guskey model was centered on the development of six overarching questions, each of which is aligned with the five levels of the Guskey model while focusing on criteria and areas of interest that are unique to the TIS program (Figure 4.1).

Figure 4.1: Guiding Evaluation Questions

QUESTION CODE	EVALUATION QUESTION	GUSKEY LEVEL
EQ1	To what extent is the training that is provided to participating TISs of adequate quality, relevance, and usefulness?	1. Participants’ Reactions
EQ2	To what extent does the TIS program build the capacity of participating TISs to plan and facilitate (a) teaching and learning, (b) information access and delivery, and (c) program administration?	2. Participants’ Learning
EQ3	To what extent do TISs encounter barriers to successful program implementation (e.g., financial, temporal, relational, etc.)?	3. Organization Support and Change
EQ4	To what extent is the level of technology integration in TIS schools positively impacted through participation in the program?	4. Participants’ Use of New Knowledge and Skills
EQ5	In what ways have school administrators and teachers leveraged the TIS and the resources provided by the TIS?	Levels 4 and 5
EQ6	What impact has the TIS program had on students’ technology literacy in participating schools?	5. Student Learning Outcomes

Source: West Virginia Department of Education¹⁰³

WVDE relies on surveys of participants to measure the program’s impact with regards to the application of learning (EQ4), and has used a mixed-methods approach with quantitative and qualitative techniques to analyze survey results and measure changes over time.

¹⁰⁰ “The West Virginia Special Education Technology Integration Specialist Program.” West Virginia Department of Education, 2012. p. 1.

<http://wvde.state.wv.us/research/reports2012/WVSpecialEducationTISProgramAdministratorsReport2012.pdf>

¹⁰¹ Ibid., p. 2.

¹⁰² Ibid.

¹⁰³ Taken verbatim from Ibid., p. 3.

Specifically, the Office of Research provided online pre- and post-training surveys to administrators (including principals, assistant principals, and teachers) at the schools where TIS participants worked.¹⁰⁴ Respondents were asked to answer pre-test questions explaining how they anticipated using specialist services once they were trained, and post-test questions retrospectively assessed how they had used specialists. Questions used a 5-point, Likert-type scale, ranging from “unlikely” to “likely.” In addition, respondents were asked via open-ended questions to write descriptive responses to the questions in addition to the ratings.¹⁰⁵ Quantitative analysis tools included descriptive statistics (such as average response rates) and tests of statistical significance. Qualitative responses to open-ended questions were coded in order to track the mention frequency of broad themes.¹⁰⁶ Copies of each survey instrument can be found in Appendix C.

Usage of the Guskey model and cohort comparisons also allowed evaluators to identify multiple areas of improvement for the Special Education system.¹⁰⁷ For example, decline in average scores between pre- and post-training on administrators’ perceived likelihood of using specialists versus their actual usage seemed to indicate that schools needed ongoing post-training support on how to use specialist services.¹⁰⁸ WDVE’s experience also suggests that regardless of the assessment tools used, evaluators must be careful to monitor administration carefully. For example, comparison of pre- and post-training data was complicated by the fact that, at some schools, different administrators took each test.¹⁰⁹

OHIO ADULT AND BASIC LITERACY EDUCATION

The Ohio Adult Basic and Literacy Education (ABLE) program provides free services to adults seeking courses on basic skills, college, career, GED prep, and English as a second language.¹¹⁰ Its professional development evaluation framework is a core part of its pursuit of continuous improvement and draws on both the Guskey and Kirkpatrick models.¹¹¹

Figure 4.2 provides an overview of ABLE’s professional development evaluation framework. ABLE emphasizes that each subsequent level of analysis builds on the prior one, such that participant satisfaction is the foundation of the evaluation pyramid, and an understanding of impact requires having a comprehensive understanding of participant satisfaction, learning, and behavior.¹¹²

¹⁰⁴ Ibid.

¹⁰⁵ Ibid., p. 4.

¹⁰⁶ Ibid., pp 4 and 5.

¹⁰⁷ Ibid., p. 16

¹⁰⁸ Ibid., pp. 4,7 and 16

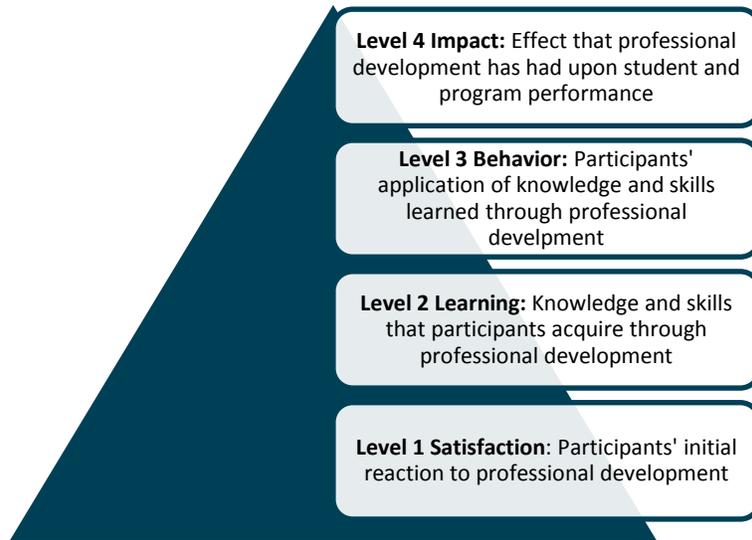
¹⁰⁹ Ibid., p. 5

¹¹⁰ “Ohio Adult Basic and Literacy Education.” ABLE. <http://www.ohioable.org/>

¹¹¹ Mullins, D., T. Lepicki, and A. Glandon. “A Professional development Evaluation Framework for the Ohio ABLE System.” Ohio State University Center on Education and Training for Employment, December 2010. pp. 1 and 2. http://uso.edu/network/workforce/able/reference/development/PD_Eval_Framework_Report.pdf

¹¹² Ibid., p. 3

Figure 4.2: Ohio ABLÉ Evaluation Framework



Source: Mullins, D., T. Lepicki, and A. Glandon ¹¹³

For each level in the evaluation framework, ABLÉ provides a brief description of what this level entails; an overview of its purpose and the intended usage of the data collected; and an overview of the data collection methodology.¹¹⁴ ABLÉ further emphasizes that not every professional development activity can be analyzed at all four levels. While most or all activities can generally be assessed at the basic level of participant satisfaction, it might be more challenging to assess how behavior or achievement was changed as a result of program participation.¹¹⁵

For Levels 1 and 2, ABLÉ primarily relies on the same core survey instrument. In fact, there is an optional component that can be added or removed from the survey depending on whether the evaluation is focused on gauging participants' reactions or learning. This optional component includes the assessment of the Knowledge, Skills, and Attitudes (KSA) of interest to evaluators.¹¹⁶ For participant learning, ABLÉ also encourages local programs to develop an additional assessment instrument based on pre-established learning objectives and to create rubrics that can analyze additional qualitative indicators such as reflective papers and lesson plans.¹¹⁷

For Levels 3 and 4, ABLÉ does not provide pre-established survey instruments because of the broad array of trainings local programs offer, and instead emphasizes the importance of

¹¹³ Adapted from *Ibid.*, p. 3

¹¹⁴ *Ibid.*, p. 5

¹¹⁵ *Ibid.*, p. 4

¹¹⁶ *Ibid.*, pp. 1 and 6.

¹¹⁷ *Ibid.*, p. 8.

a general cohesive evaluation process. However, Figure 4.3 presents further guidance regarding the measurement and evaluation of participant behavior.

Figure 4.3: Level 3 Evaluation Process

DEFINE THE BEHAVIOR OBJECTIVES OF THE TRAINING
The provider of the professional development should define which behaviors the training is attempting to increase, decrease, or otherwise modify as a result of the training.
SPECIFY DIMENSIONS OF QUALITY AND QUANTITY FOR THE BEHAVIOR OBJECTIVES
In defining the behavior objectives, the provider should define the criteria for measures of desirable behavior including the frequency with which the behaviors should take place.
DETERMINE TIME DURATION BETWEEN TRAINING AND EVALUATION
Since participants need time to plan and reflect on how to implement knowledge and skills gained through a training, providers will need to decide how long to wait before evaluating the success or failure of implementation. Depending on the complexity of the behavior objectives, this delay could range from one week to three months. Providers may also consider conducting a second round of evaluation within six months of the training as a follow-up to the initial evaluation.
DETERMINE METHODS OF EVALUATION
The training provider will need to decide upon one or more evaluation methods to utilize for evaluating changes in behavior. Possible methods include: <ul style="list-style-type: none"> ▪ Onsite observations of participants ▪ Written descriptions of implementation process by participants (reflective journals, portfolios, etc.) ▪ Follow-up interviews with participants ▪ Self-reporting evaluations on implementation

Source: Mullins, D., T. Lepicki, and A. Glandon¹¹⁸

Regarding Level 4, ABLE notes that “impact evaluations are the most complex and difficult to implement.”¹¹⁹ ABLE observes that this level of analysis generally relies on existing data sources and recommends the following sources as resources:

- ABLELink data (student records)
- Record of Accomplishment section of the Individual Professional Development Plan
- Program Professional Development Plan
- Local Program Desk Review
- Local Program Data Quality Checklist (staff training)¹²⁰

ABLE also maintains that longitudinal tracking of changes that compare outcomes for trained participants versus those who have yet to be trained is ideal, but may require certain logistical conditions, such as staggered roll-outs and multi-site or multi-year initiatives. Regardless of the type of data used, the outcomes and the measurement criteria for the data sources should be defined prior to the training.¹²¹

¹¹⁸ Ibid., p. 10.

¹¹⁹ Ibid., p. 12.

¹²⁰ Ibid., p. 13.

¹²¹ Ibid.

APPENDIX A

OHIO ABLE EVALUATION FRAMEWORK: LEVEL 1

Level 1: Satisfaction Instrument

Below is an excerpt from the current standardized evaluation instrument used in all Ohio ABLE single session trainings to collect participant satisfaction data.

1. My goal for attending this session is:

2. On this topic, I consider myself (choose one):

An Expert	Skilled/Knowledgeable	A Novice
③	②	①

SESSION CONTENT:

In regards to this session, the content presented...

	Significantly	Generally	Somewhat	Very Little	Don't Know
3. is USEFUL to me.....	④	③	②	①	①
4. is APPLICABLE to my job.....	④	③	②	①	①
5. has CHANGED my THINKING.....	④	③	②	①	①
6. has REINFORCED my THINKING.....	④	③	②	①	①

7. List at least one thing you learned today that you will use in your classroom/program.

Concerning the content of the session you attended, how much have each of the following INCREASED?

8. KNOWLEDGE of the content presented.....	④	③	②	①	①
9. CONFIDENCE that you can apply the knowledge to your job.....	④	③	②	①	①
10. MOTIVATION to implement the content/techniques presented.....	④	③	②	①	①

Source: Ohio ABLE¹²²

¹²² Mullins et al., Op. Cit., p. 20

OHIO ABLE EVALUATION FRAMEWORK: OPTIONAL LEVEL 2 COMPONENT

KNOWLEDGE, SKILLS, AND ATTITUDE (Before and after this training)

11. [INSERT KSA]	Before	④	③	②	①	①
	After	④	③	②	①	①
12. [INSERT KSA]	Before	④	③	②	①	①
	After	④	③	②	①	①
13. [INSERT KSA]	Before	④	③	②	①	①
	After	④	③	②	①	①
14. [INSERT KSA]	Before	④	③	②	①	①
	After	④	③	②	①	①
15. [INSERT KSA]	Before	④	③	②	①	①
	After	④	③	②	①	①

Source: Ohio ABLE¹²³

Note: This portion of the form would be customized to reflect the specific Knowledge, Skills, and Attitude sought by the professional development program.

OHIO ABLE EVALUATION FRAMEWORK: LEVEL 3¹²⁴

This appendix contains sample questions that can inform the collection of data to measure changes in behavior as a result of professional development. These samples are intended to provide a basic understanding of behavior evaluation. Professional development providers would customize the questions and consider a variety of methods for collecting the data (e.g., interview, observation, questionnaire, reflective journal).

Information for Implementation

1. List at least one thing you have implemented in your classroom/program from the training.
2. Explain one “take away” from the training that has stuck with you.

Description of Implementation

3. Since the training, how have you used the strategies in your classroom/program?
4. Comparing the training to your current practice, how has your practice improved because of the training?
5. What have you done differently in your practice as a result of the training?
6. How do you vary your implementation of what you learned in the training in order to accommodate your classroom?
7. Reflecting on your current practices, are they:

¹²³ Ibid.

¹²⁴ Mullins et al. Op. Cit., p. 22.

- directly influenced by what you learned in the training
- influenced by participating in the training
- as a result of another source (explain)

Degree of Implementation

8. To what extent has the information in the training changed your classroom routine

9. How regularly are you using the techniques presented in the training in your program?

10. To what extent have you integrated the strategies from the training into your work?

APPENDIX B

GADSEN ELEMENTARY SCHOOL DISTRICT #32: ORGANIZATIONAL CONTEXT INDICATORS

The principal and other leaders do the following.

<i>Desired outcome</i>	<i>LEVEL I</i>	<i>LEVEL II</i>	<i>LEVEL III</i>	<i>LEVEL IV</i>
Create atmosphere or context for change	Schedule time and place for staff reflection and collaborative work Provide learning environment Develop culture of learning Develop staff's skills of <ul style="list-style-type: none"> • collaboration • modes of conversation • conflict management • decision-making model Nurture leadership team skills Activate leadership teams for learning Monitor to ensure time is used well	Schedule time and place for staff reflection and collaborative work Provide learning environment Develop culture of learning Develop staff's skills of <ul style="list-style-type: none"> • collaboration • modes of conversation • conflict management • decision-making model, Activate leadership teams for learning	Schedule time and place for staff reflection and collaborative work Provide learning environment Develop staff's skills of <ul style="list-style-type: none"> • collaboration • modes of conversation • conflict management • decision-making models 	Schedule time and place for staff reflection and collaborative work Provide learning environment
Develop and communicate a shared vision	Identify purpose or school mission Define values and staff beliefs Engage staff in studying data to identify needs for improvement Study and select new programs or practices to address the priority need for improvement Create an innovation configuration that represents and communicates the new practice, the vision of change Keep the vision visible Revisit the vision periodically	Identify purpose or school mission Engage staff in studying data to identify needs for improvement Select new programs or practices to address the priority need for improvement Create an innovation configuration that represents and communicates the new practice, the vision of change	Engage staff in studying data to identify needs for improvement Select new programs and practices to address the priority need for improvement	Engage staff in studying data to identify needs for improvement Adopt new programs and practices to address the need for improvement

(Continued)

<i>Desired outcome</i>	<i>LEVEL I</i>	<i>LEVEL II</i>	<i>LEVEL III</i>	<i>LEVEL IV</i>
Plan and provide resources	Gather staff information (stages of concern, level of understanding, innovation configuration) and relevant data Use six strategies to develop an implementation plan that will achieve the vision Identify resources needed and plan to access them: <ul style="list-style-type: none"> • currently available • needed as reflected in the vision Establish timelines	Gather staff information (stages of concern, level of understanding, innovation configuration) and relevant data Use six strategies to develop an implementation plan that will achieve the vision Identify resources needed and plan to access them: <ul style="list-style-type: none"> • currently available • needed as reflected in the vision 	Use six strategies to develop an implementation plan that will achieve the vision Identify resources needed and plan to access them: <ul style="list-style-type: none"> • currently available • needed as reflected in the vision 	Identify resources needed and plan to access them: <ul style="list-style-type: none"> • currently available • needed as reflected in the vision
Invest in professional development	Gather and analyze student data (AIMS, mandated and district assessments) and teacher data (stages of concern, level of understanding, innovation configuration) Use staff and student data to create adult learning activities Create vision-driven action plan for professional development Arrange for, schedule, and deliver adult learning activities Establish timelines	Gather and analyze student data (AIMS, mandated and district assessments) and teacher data (stages of concern, level of understanding, innovation configuration) Use staff and student data to create adult learning activities Create vision-driven professional development plan	Gather and analyze student data (AIMS, mandated and district assessments) and teacher data (stages of concern, level of understanding, innovation configuration) Create professional development plan	Create professional development plan

<p>Check for progress</p>	<p>Gather staff information</p> <ul style="list-style-type: none"> • stages of concern • level of understanding • innovation configuration <p>Include staff in interpreting data and determining needs</p> <p>Develop a culture of continuous assessment</p> <p>Celebrate small and large successes publicly or privately</p> <p>Establish timelines</p>	<p>Gather staff information</p> <ul style="list-style-type: none"> • stages of concern • level of understanding • innovation configuration <p>Include staff in interpreting data and determining needs</p> <p>Develop a culture of continuous assessment</p>	<p>Gather staff information</p> <ul style="list-style-type: none"> • stages of concern • level of understanding • innovation configuration <p>Ask staff what they need for implementation of new practices</p>	<p>Ask staff what they need for implementation of new practices</p>
<p>Provide assistance</p>	<p>Schedule needed professional development</p> <ul style="list-style-type: none"> • large-group • small-group • individuals <p>Provide coaches or mentors</p> <p>Review time and activities for collaborative work</p> <p>Inventory resources, restock, or share</p> <p>Revisit action plan and revise as needed</p> <p>Celebrate small and large successes publicly or privately</p>	<p>Schedule needed professional development</p> <ul style="list-style-type: none"> • large-group • small-group • individuals <p>Provide coaches or mentors</p> <p>Review time and activities for collaborative work</p> <p>Inventory resources and restock</p>	<p>Schedule needed professional development</p> <ul style="list-style-type: none"> • large-group • small-group • individuals <p>Review time and activities for collaborative work</p> <p>Inventory resources and restock</p>	<p>Schedule needed professional development</p> <ul style="list-style-type: none"> • large-group

Source: Copyright © 2007 Gadsden Elementary School District #32. Reprinted with permission from the leadership team at Gadsden Elementary School District #32, San Luis, AZ; created in collaboration with Wilda Storm and Shirley Hord.

Source: Joellen Killion¹²⁵

¹²⁵ Killion, Op. Cit., pp. 187-189.

APPENDIX C

WVDE SPECIAL EDUCATION TIS ADMINISTRATOR PRE SURVEY¹²⁶

This survey is intended to help you to determine the most effective ways to utilize the SE TIS in your school. Please note that you may receive a second survey near the end of the school year to help WVDE better understand how school administrators are leveraging the resources provided to schools through the SE TIS program.

Section I: About You

In which county is your school located?

Please indicate the name of your school.

What is your role within your school?

- Principal
- Assistant Principal
- Other

Before receiving this survey, I was aware that my school would have a SE TIS for the 2011-12 academic year.

- Yes
- No

Section II: Your plans to use the SE TIS

Please indicate how likely it is that you will use the SE TIS for each of the following purposes.

I plan to ask my SE TIS to share what he/she has learned by leading standards-based professional development for the other teachers in my school.

Unlikely 1 2 3 4 5 Likely

I plan to ask my SE TIS to model the integration of technology for their co-teachers and others within the school.

Unlikely 1 2 3 4 5 Likely

I anticipate asking my SE TIS to assist me in developing the school's strategic plan with regard to information and technology needs.

Unlikely 1 2 3 4 5 Likely

I expect that my SE TIS will assist his/her co-teachers in customizing available digital resources and tools such as West Virginia Writes (formally Writing Roadmap), TechSteps, and Acuity to personalize learning for students.

Unlikely 1 2 3 4 5 Likely

¹²⁶ West Virginia Department of Education, Op. Cit., pp. 19-22

I will request that the SE TIS work with teachers to identify digital resources and tools that effectively integrate technology into their current curriculum.

Unlikely 1 2 3 4 5 Likely

I will ask the TIS to conduct analyses of student data and engage in action research to help me understand the impact of technology integration in my school.

Unlikely 1 2 3 4 5 Likely

Please articulate what you believe the role of the SE TIS should be in your school.

What outcomes do you expect as a result of having a SE TIS in your school?

Please provide any additional comments you may have about the SE TIS program.

WVDE SPECIAL EDUCATION TIS ADMINISTRATOR POST SURVEY

This survey is intended to help WVDE better understand how school administrators are leveraging the resources provided to schools through the SE TIS program. Thank you for your cooperation and support of the SE TIS program.

Section I: About You

In which county is your school located?

Please indicate the name of your school.

What is your role within your school?

- Principal
- Assistant Principal
- Other

Before receiving this survey, I was aware that my school would have a SE TIS for the 2011-12 academic year.

- Yes
- No

Section II: Use of the SE TIS

Please indicate your use of the SE TIS for each of the following purposes.

My SE TIS shared what he/she learned by leading standards-based professional development for the other teachers in my school.

Seldom 1 2 3 4 5 Always

My SE TIS modeled the integration of technology for their co-teachers and others within the school.

Seldom 1 2 3 4 5 Always

My SE TIS assisted me in developing the school's strategic plan with regard to information and technology needs.

Seldom 1 2 3 4 5 Always

My SE TIS assisted his/her co-teachers in customizing available digital resources and tools such as West Virginia Writes (formally Writing Roadmap), TechSteps, and Acuity to personalize learning for students.

Seldom 1 2 3 4 5 Always

I requested that the SE TIS work with teachers to identify digital resources and tools that effectively integrate technology into their current curriculum.

Seldom 1 2 3 4 5 Always

I asked the TIS to conduct analyses of student data and engage in action research to help me understand the impact of technology integration in my school.

Seldom 1 2 3 4 5 Always

Please articulate what you believe the role of the SE TIS should be in your school.

What outcomes did you observe as a result of having a SE TIS in your school?

Please provide any additional comments you may have about the SE TIS program.

PROJECT EVALUATION FORM

Hanover Research is committed to providing a work product that meets or exceeds partner expectations. In keeping with that goal, we would like to hear your opinions regarding our reports. Feedback is critically important and serves as the strongest mechanism by which we tailor our research to your organization. When you have had a chance to evaluate this report, please take a moment to fill out the following questionnaire.

<http://www.hanoverresearch.com/evaluation/index.php>

CAVEAT

The publisher and authors have used their best efforts in preparing this brief. The publisher and authors make no representations or warranties with respect to the accuracy or completeness of the contents of this brief and specifically disclaim any implied warranties of fitness for a particular purpose. There are no warranties that extend beyond the descriptions contained in this paragraph. No warranty may be created or extended by representatives of Hanover Research or its marketing materials. The accuracy and completeness of the information provided herein and the opinions stated herein are not guaranteed or warranted to produce any particular results, and the advice and strategies contained herein may not be suitable for every partner. Neither the publisher nor the authors shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages. Moreover, Hanover Research is not engaged in rendering legal, accounting, or other professional services. Partners requiring such services are advised to consult an appropriate professional.



4401 Wilson Boulevard, Suite 400

Arlington, VA 22203

P 202.559.0500 F 866.808.6585

www.hanoverresearch.com